

## **Análisis exploratorio de juicios morales en la discusión de dilemas. Profundizando en la formación ética canalizada mediante la tecnología<sup>1</sup>**

Josefa Cerda Maureira, Camilo Carvajal Reyes,  
Pablo Ramírez Rivas, Eduardo Hurtado Mila<sup>2</sup>

La Unidad de Ética (UE) del Área de Estudios Transversales en Humanidades para las Ingenierías y las Ciencias (ETHICS), perteneciente a la Escuela de Ingeniería y Ciencias (EIC) de la Facultad de Ciencias Físicas y Matemáticas (FCFM), tiene como propósito formar en competencias éticas al estudiantado. Una de las actividades que realiza con ese fin es la discusión en torno a dilemas éticos. El caso por estudiar se denomina Caso Adela y para ello se utiliza la aplicación *EthicApp*. El gran volumen de datos dificulta el análisis. Este trabajo aborda el uso de modelos de aprendizaje de máquina supervisados y no-supervisados (*Naive-Bayes* y *BERTopic*, respectivamente) para modelar la estructura textual de las respuestas y con esto apoyar el análisis que puedan hacer los equipos docentes. Pese al potencial de mejora, la metodología ofrece una visión general de las respuestas y conceptos utilizados, lo que aportará en

---

<sup>1</sup> Agradecemos a Gustavo Zurita (FEN, UChile) y a Claudio Álvarez (UAndes) por facilitarnos el uso de la aplicación web *EthicApp*, además de su constante colaboración y paciencia para la realización de las actividades formativas. Agradecemos también a Claudia Rodríguez, Coordinadora de ETHICS, por acompañar y apoyar el trabajo de la Unidad de Ética. Finalmente, un agradecimiento a Ignacio González, coordinador del Área de Ingeniería e Innovación, por su compromiso y tiempo para desarrollar la competencia ética en los cursos de plan común.

<sup>2</sup> Josefa Cerda Maureira, Universidad de Chile, [jozefacerda@uchile.cl](mailto:jozefacerda@uchile.cl); Camilo Carvajal Reyes, Universidad de Chile, [ccarvajal@dim.uchile.cl](mailto:ccarvajal@dim.uchile.cl); Pablo Ramírez Rivas, Universidad de Chile, [pabramirez@uchile.cl](mailto:pabramirez@uchile.cl); Eduardo Hurtado Mila, Universidad de Chile, [eduardo.hurtado@ug.uchile.cl](mailto:eduardo.hurtado@ug.uchile.cl).

la evaluación de la competencia ética de estudiantes de la FCFM. Los resultados obtenidos permiten concluir que los y las estudiantes no incluyen explícitamente los principios éticos en su justificación, pudiendo esto implicar que se encuentran en un nivel básico de la competencia; además, cuando la responsabilidad se enfrenta con el respeto (como ocurre en el Caso Adela), el estudiantado se inclinará por el principio de responsabilidad.

## Introducción

La formación de la ética (Cortina, 2000) plantea una serie de interrogantes y desafíos: ¿cómo puede enseñarse un tipo de saber de naturaleza práctica en contextos masivos como el de la FCFM? ¿Cómo se evalúa y se plantea en indicadores de evaluación un desempeño que tiene lugar en tan diversos contextos? ¿Puede o debe la formación en ética hacerse cargo de todas las actuaciones de quienes son sujetos de enseñanza? ¿Cómo se forma «el carácter» de las personas hoy en la educación superior?, ¿se puede hacer? ¿Cómo pueden abordarse estas dificultades en el diseño, implementación y evaluación de un modelo de competencia ética exitoso en cuanto al aprendizaje de los y las estudiantes? La EIC ha tomado parte en estos desafíos y creó hace dos años la UE al interior de ETHICS con el propósito de hacer frente a estas interrogantes con personas formadas y especializadas en el tema.

En el XXXIV Congreso de la Sociedad Chilena de Educación en Ingeniería, la Unidad de Ética presentó la ponencia «La formación ética canalizada mediante la tecnología: experiencia y resultados preliminares del uso de la herramienta web *EthicApp*» (cfr. Ramírez, P., Guerrero, S., Cerda, J., Ross, J. y Flores, G., 2022) en la cual se aborda un caso ético denominado «Caso Julieta» en una actividad de formación en una asignatura obligatoria de la carrera de ingeniería. Los principales resultados fueron: (1) el impacto que tiene el diálogo y la confrontación de ideas en el enriquecimiento de los propios juicios, puntos de vista y preferencias morales, (2) luego de la fase grupal del trabajo del caso, los y las estudiantes tienden a tomar posturas más moderadas que extremas, (3) el 60% de los y las

estudiantes mantienen la postura grupal de la etapa previa a la etapa individual final y (4) la postura de cometer un acto deshonesto está basada en argumentos de juicio moral que integran elementos que van más allá del beneficio personal y que consideran la relación y corresponsabilidad con otros.

Este artículo pretende ser una continuación y profundización de la ponencia ya mencionada por medio de un análisis exploratorio más profundo y acabado de las justificaciones escritas. Debido a la naturaleza masiva de la formación en ingeniería, la cantidad de texto a analizar representa un desafío para la evaluación; por ello se utilizarán herramientas de procesamiento de lenguaje natural, en particular, modelos de aprendizaje de máquina supervisados y no-supervisados para modelar la estructura textual de las respuestas. Esto permitirá comprender con más detalle las elecciones de los y las estudiantes, la identificación de juicios morales y los principios éticos o elementos que consideran para discernir por una determinada postura.

## **El contexto**

La Escuela de Ingeniería y Ciencias recibe cada año más de mil estudiantes en su programa de Plan Común, que dura cuatro semestres y es un paso obligatorio para elegir una especialidad o licenciatura. Durante los primeros tres semestres del programa de Plan Común, los y las estudiantes deben rendir los siguientes cursos en el orden señalado: Desafíos de innovación, Proyecto en innovación en Ingeniería y Ciencias y Módulo interdisciplinario. Estos cursos son parte de la línea de formación impartida por una unidad de la Escuela denominada Área de Ingeniería e Innovación que, entre otros objetivos, busca desarrollar las competencias transversales de innovación, trabajo en equipo, comunicación oral y escrita y compromiso ético.

Cada uno de estos cursos tiene diez secciones de cien estudiantes cada una, cantidad que representa un desafío tanto para los equipos docentes, en tanto formadores en competencias

específicas de las ingenierías y las ciencias, así como para las unidades a cargo de la formación y evaluación de las competencias transversales o genéricas. Por ello, la Unidad de Ética, a cargo del desarrollo de la competencia de compromiso ético, ha buscado distintas estrategias "algunas de ellas serán mencionadas ahora" para cumplir con el propósito de aprendizaje para lo cual debe levantar evidencias del desempeño de los y las estudiantes.

La competencia transversal de compromiso ético está dividida en dos momentos de acuerdo con el ciclo de formación del que se trate. En el caso del curso en que se aplicará el caso en cuestión "en adelante Caso Adela", corresponde la competencia transversal en su momento inicial: «Reflexionar sobre el propio actuar y sus consecuencias, en el marco de la honestidad, la responsabilidad y el respeto, buscando la excelencia y rigurosidad en su proceder en contextos académicos, en las relaciones interpersonales y con su entorno». Esta competencia se manifiesta a partir de una serie de indicadores de evaluación que se relacionan con los principios de responsabilidad, respeto e integridad en contextos profesionales.

A continuación, se comparte una síntesis del caso Adela y sus dilemas morales distribuidos en etapas según diferenciales semánticos<sup>3</sup>:

*En Chile, la deficiencia de vitamina D es un problema serio tanto en adultos mayores como en niños. Para abordar esta preocupante situación, un grupo de profesionales creó una startup para buscar una solución. Investigando, encontraron una fruta ancestral de las comunidades diaguitas con alta concentración de vitamina D y un sabor atractivo para el consumo. Adela, una ingeniera del equipo, diseña el proceso de producción de un nuevo alimento a partir de esta fruta. Sin embargo, se enfrenta a desafíos, ya que el árbol solo crece cerca de los ríos y necesita abundante luz solar, lo que dificulta llevarlo a zonas más australes; por lo tanto, debe cultivarse en el Norte Chico, región afectada por la sequía. Además,*

---

<sup>3</sup> Para más información sobre la metodología de dilemas morales ver sección correspondiente en ponencia anterior: <https://revistasdex.uchile.cl/index.php/cdb/article/view/12153/12192>

para conservar la vitamina D durante el transporte a otras zonas, el equipo decide liofilizar la fruta y agregar conservantes. Aunque aún no tienen la obligación legal de integrar a las comunidades diaguitas en el proyecto, Adela escucha sus preocupaciones sobre cómo estos cambios afectarían sus tradiciones. Pese a que los cambios son necesarios para ayudar a quienes sufren deficiencia de vitamina D, las comunidades prefieren mantener sus prácticas tradicionales, ya que estas son parte fundamental de su identidad.

Para el análisis de este caso, se trabaja en torno a dos diferenciales semánticos<sup>4</sup> presentados a continuación: **DS1** - Considerando el uso de agua en un sector de escasez hídrica para la producción de este nuevo alimento y la potencial población beneficiada, Adela debería priorizar... : (1) Preservar el recurso natural escaso o (6) Beneficiar la salud de ancianos y niños; **DS2** - Considerando la futura población beneficiada por este nuevo alimento y los efectos generados a las tradiciones identitarias del pueblo originario del caso, Adela debería priorizar...: (1) Producir el alimento contra déficit vitamínico o (6) Resguardar las tradiciones identitarias. En el modelo de la Figura 1. se esquematiza lo anterior.

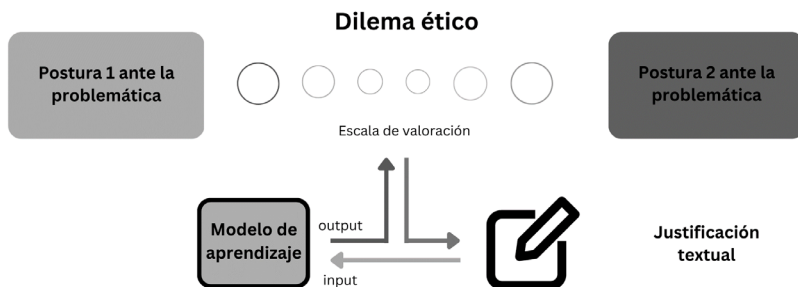


Figura 1. Esquema de modelo para predecir posturas de estudiantes.

<sup>4</sup> Cada diferencial semántico tiene seis opciones dentro de las cuales los y las estudiantes pueden escoger según el grado de acercamiento que tengan con determinada postura. El número de opciones busca evitar posturas neutras.

El caso en cuestión ha tenido tres aplicaciones supervisadas por la UE de ETHICS. En el semestre de primavera del año 2021 se aplica en el tercer curso (CD2201) de la línea de cursos ya mencionados; en el semestre de primavera del año 2022, se aplica en el segundo curso (CD1201) y, por último, en el semestre de otoño del año 2023 se aplica en el primer curso (CD1100). En la Tabla 1 se muestran las realizaciones del caso, las etapas de cada una, el número de personas analizadas y el porcentaje de realización de la actividad. La modalidad de aplicación del caso se ejecuta en cuatro etapas: Etapa 0 (E0), cuya realización es previa al día de la actividad sincrónica que tiene como objetivo que los y las estudiantes creen sus cuentas de usuario/a en *EthicApp*; Etapa Individual 1 (EI-1) en la cual se lee el caso en clases y se responden los dos diferenciales semánticos; Etapa Grupal en la cual los y las estudiantes son asignados a grupos con criterio de heterogeneidad en sus respuestas y discuten grupalmente en torno a los dos diferenciales; por último, la Etapa Individual 2 (EI-2) en la que se vuelve a responder individualmente luego de la etapa grupal.

Tabla 1  
*Resumen de las realizaciones del caso Adela.*

Año	Curso	Etapas	Número de personas analizadas	Porcentaje realización
2023-1	CD1100	E0, EI-1, EG, EI-2	811 de 866	94%
2022-2	CD1201	E0, EI-1, EG, EI-2	591 de 716	83%
2021-2	CD2201	EI-1, EG, EI-2	160 de 213	75%

Debido a los límites de extensión del presente artículo, en adelante se analizará únicamente el diferencial semántico 2 y su comportamiento en la EI-1, EG y EI-2.

## Metodología

Para remediar la limitación de la gran magnitud de datos para el análisis de las justificaciones, se propone el uso de modelos para

procesamiento de texto que se usan en el contexto de la ciencia de datos y la inteligencia artificial. Estos modelos cuentan con una base estadística y permiten realizar predicciones y/o separaciones de los datos que se desean modelar, siendo estos los textos de estudiantes en este caso. El método que se propone consiste en utilizar estos modelos y ajustarlos a la tarea de predecir la valoración frente al dilema ético. Esto permitirá, con un análisis a posteriori, investigar qué elementos semánticos son los más relevantes a la hora de tomar una u otra postura, que es el objetivo final.

En el contexto del procesamiento de lenguaje natural se han estudiado ya las respuestas a preguntas o cuestiones morales. Muchos de estos trabajos toman como base la teoría de fundamentos morales (TMF) propuesta por Graham et al. (2013), siendo Garten, Boghrati, Hoover, Johnson y Dehghani (2016) y Xie, Hirst y Xu (2020) dos ejemplos. En el primer trabajo se busca la predicción/identificación de retóricas morales y, en el segundo, se evalúan modelos que también predicen e identifican retóricas morales. Por otro lado, Kennedy et al. (2021) buscan la predicción de preocupaciones morales propias de un individuo usando evidencias de lenguaje moral que escriben usuarios y usuarias en redes sociales. Se utilizaron distintas técnicas de procesamiento de lenguaje para predecir los puntajes obtenidos por usuarios y usuarias, para cada una de las dimensiones morales planteadas en la TMF. Se destaca la variedad de métodos testeados para vectorizar texto, incluyendo *latent dirichlet allocation (LDA)* (Blei, Ng y Jordan, 2003), representaciones de palabras distribucionales (*word embeddings*), conteo de ocurrencias de conceptos relacionados a TMF y *BERTopic* (Devlin, Chang, Lee y Toutanova, 2019), que es un modelo de lenguaje profundo. Es este último el que obtiene mejores resultados. Finalmente, tanto conteos de ocurrencias de conceptos como *LDA* se usaron para interpretar qué elementos lingüísticos específicos explicaban cada fundamento por separado.

Los trabajos anteriores muestran la alta variedad de formatos en los cuales se ha evaluado la presencia y grado de categorías morales, así como también la capacidad de distintos modelos de procesamiento de lenguaje natural para modelarlas. No obstante, ninguno de los artículos estudiados enfrenta un desafío tan específico

como el de este trabajo. Por ello, la modalidad consistirá en una exploración de características básicas de las justificaciones textuales "que se considerará el único elemento a utilizar", para luego usar modelos predictivos que modelen el texto. Por su parte, el valor por predecir corresponderá a la opción ante la problemática en cuestión. Este proceder busca obtener una fotografía global de las justificaciones de estudiantes, condicionadas a la preferencia escogida.

Cabe señalar que tener un modelo que prediga la postura (concretizada por la opción escogida) dada en el texto no es de interés *per se*. La ventaja es que, a través de tal herramienta, se identifiquen los elementos semánticos que justifiquen los argumentos dados por los y las estudiantes. Es por esto que se preferirán modelos que tengan capacidad interpretativa, porque de esta manera se sabe qué elementos del *input* (secuencias de palabras en este caso) hacen que el resultado sea uno y no otro. Estos elementos (dado un modelo bien ajustado a los datos) mostrarán entonces el comportamiento general de las justificaciones, facilitando así el análisis de la gran cantidad de textos que se generan en cada actividad.

El modelo supervisado *Naive-Bayes*, que naturalmente asigna probabilidades por postura a cada palabra, es un modelo de clasificación que será testeado separando tanto las seis posturas y casos independientes, así como agrupando las opciones en dos. Al analizar las probabilidades se pueden identificar aquellos elementos (palabras o tópicos, según la vectorización) que contribuyan más al resultado final del algoritmo, obteniendo la capacidad interpretativa que se señalaba. Se concluye con la familia de modelos no-supervisados llamada Modelamiento de Tópicos, para el cual se probaron tanto el modelo *LDA* como *BERTopic* (Grootendorst, 2022), basado en redes neuronales. Estos modelos también permiten asociar palabras y textos completos a probabilidades de clase, pero esta vez las clases corresponderán a los tópicos encontrados.

## Resultados

En la Figura 2, el gráfico agregado de las tres realizaciones del caso Adela muestra que las opciones 1, 2 y 3 son predominantes a lo



largo de la actividad, siendo en la EG donde se agudiza la diferencia, por tanto, hay una clara concentración de la muestra de estudiantes en una postura, oscilando entre el 64% y el 71% del total. En particular, la mayor cantidad de estudiantes se concentra en el valor 2 en la EI-1 (29%) y en el valor 3 para la etapa grupal e individual 2 (42% y 30%, respectivamente). Destaca que en la EI-1 el 21% escoge la opción 3, mientras que en la EG esta preferencia aumenta al 42%, para luego, en la etapa 3, disminuir a 30%. Este comportamiento señala que en la etapa grupal las posturas se moderan, posiblemente por el contraste de ideas que se generan entre quienes conforman el grupo.

El examen exploratorio básico incluyó un análisis de la extensión de las respuestas (número de palabras). El promedio de palabras es inferior a 50 para el caso estudiado. Las opciones (1) y (6) son menos extensas que el promedio en comparación con las opciones más moderadas (2, 3, 4 y 5). Este fenómeno puede responder al hecho de que una persona absolutamente convencida de su postura vea menos necesidad de explicarla. De cualquier modo, la diferencia es sutil.

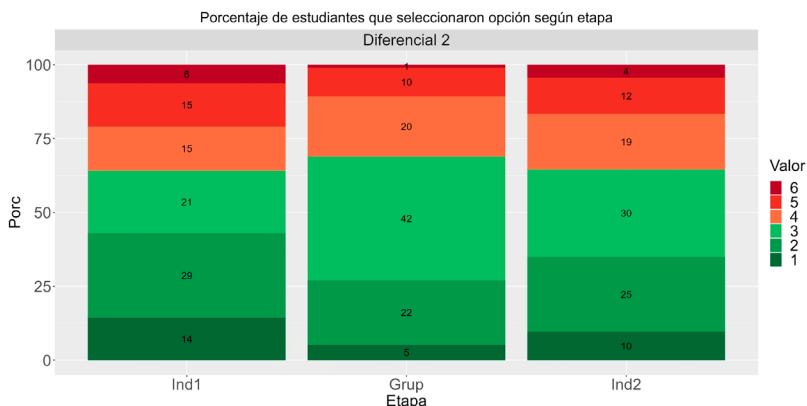


Figura 2. Porcentaje de estudiantes que seleccionaron opción según etapa (EI-1, EG, EI-2), Caso Adela en las tres realizaciones.

Por otro lado, se analizan las palabras más comunes condicionado cada postura (ver Figura 3). Se puede notar que algunas palabras utilizadas por los y las estudiantes para justificar son las mismas que las palabras de cada postura para el diferencial. Por ejemplo, se aprecia que para apoyar la producción del alimento se evocan conceptos como «salud» y «personas», así como también acciones como «producir» y «priorizar», este último reflejando que estudiantes que asumen esa postura probablemente priorizan el bien común y/o salud de niños y ancianos en su decisión. En contraste, sustantivos como «cultura» y «pueblos» (en plural) aparecen al escoger resguardar las tradiciones. Asimismo, expresiones como «resguardar», «respetar» y «debería» son usadas con el mismo propósito. Este análisis también permite identificar palabras ampliamente repetidas sin importar la postura del estudiante, como es el caso de «tradiciones». Adicionalmente, el concepto «acuerdo» aparece solo en las opciones (3) y (4), que son las más moderadas del espectro de posturas.

En la nube de palabras (Figura 4) se muestra una visualización de la frecuencia de palabras por postura (la postura de la izquierda corresponde a las opciones 1, 2 y 3 y la de la derecha a las opciones 4, 5 y 6) y se agregan, además, pares de palabras consecutivas. Una mayor frecuencia se traduce en un mayor tamaño de la palabra en la nube. Se distingue que, en la nube de la izquierda, que corresponde a la postura de *producir alimento contra déficit vitamínico*, aparece la palabra «beneficio», posiblemente para aportar mayor comprensión a la elección. Además, tanto en la nube de palabras de la izquierda como en la de la derecha, que corresponde a la postura de *resguardar tradiciones identitarias*, se repite la palabra «acuerdo». Las otras palabras y frases no aportan información adicional a la comprensión de las justificaciones porque aparecen en la descripción del caso, o bien, en la construcción del diferencial semántico.

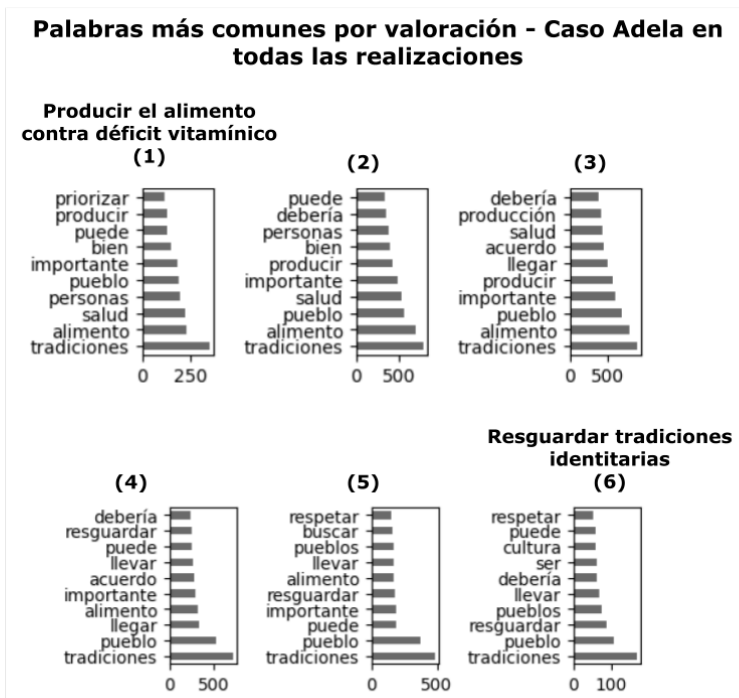


Figura 3. Frecuencias de posturas por etapa en caso Adela.

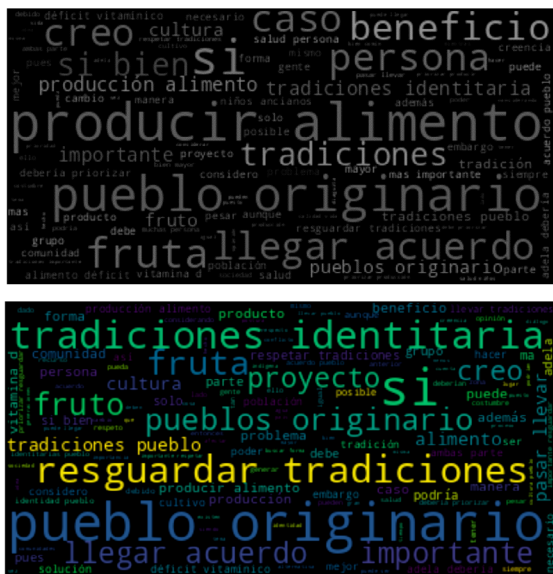


Figura 4. Nube de palabras por postura para el caso Adela.

## Resultados de modelos estadísticos

Dado que predecir de manera exacta la opción (de 1 hasta 6) es una tarea difícil, se considera predecir la postura binaria (agrupando las opciones (1), (2) y (3), por un lado, y las opciones (4), (5) y (6), por el otro). Con esta estrategia la exactitud del modelo sube de 0.43 a 0.79 (mientras más alta, mejor es la predicción), pero además permite observar la importancia de palabras para cada postura.

Se incluyen a continuación aquellas palabras que tienen más probabilidad para cada postura, como se ha explicado en el párrafo anterior. La Tabla 2 muestra conceptos para los cuales se identificó una probabilidad alta para cada una de las posturas tomando como referencia el modelo *Naive-Bayes*. La presencia de estos conceptos

en el texto se relaciona con las posturas en cuestión, lo cual pone en evidencia los elementos semánticos más recurrentes a la hora de escribir las justificaciones. Es posible llevar a cabo un análisis similar con los modelos de tipo lineal. De la tabla destaca que para la postura de producir el alimento hay más probabilidad de que en la justificación se usen palabras como «salvar», «juego», «vidas», «tribu», «religiosas» y «riesgo»; estas palabras podrían mostrar que quienes escogen esta postura consideran que salvar vidas o las vidas que están en juego son más importantes. Igualmente, señalan a los pueblos originarios como «tribu» y para referirse a sus tradiciones las consideran de tipo «religiosas». Por otro lado, las palabras que destacan de aquellas que más probabilidad tienen de ser usadas para la postura de resguardar tradiciones identitarias son «suplementos», «sol» y «obtener»; esto podría relacionarse con que quienes escogen esta opción aluden a alternativas para abordar el déficit vitamínico vía suplementos u obtenerla a través del sol. El mismo análisis se obtiene al considerar un modelo que prediga el valor seleccionado entre 1 y 6 y puede también generalizarse a otro tipo de modelos.

Tabla 2  
*Palabras con alta probabilidad para ambas posturas con modelo Naive-Bayes*

Producir el alimento contra déficit vitamínico				Resguardar tradiciones identitarias			
Palabra	Prob	Palabra	Prob	Palabra	Prob	Palabra	Prob
1. <i>salvar</i>	0.985	6. <i>poniendo</i>	0.942	1. <i>siglos</i>	0.880	6. <i>usuario</i>	0.761
2. <i>juego</i>	0.958	7. <i>riesgo</i>	0.938	2. <i>suplementos</i>	0.863	7. <i>consentimiento</i>	0.761
3. <i>vidas</i>	0.952	8. <i>tiempos</i>	0.934	3. <i>sol</i>	0.825	8. <i>integridad</i>	0.730
4. <i>tribu</i>	0.950	9. <i>cambiarían</i>	0.934	4. <i>existen</i>	0.822	9. <i>sacar</i>	0.724
5. <i>religiosas</i>	0.948	10. <i>ayudarían</i>	0.934	5. <i>única</i>	0.770	10. <i>obtener</i>	0.723

En la Tabla 3 se muestra el resultado del modelamiento de tópicos con el método *BERTopic*. Respecto a los modelos, gracias a una inspección de los tópicos obtenidos, se concluye que el modelo *LDA* identifica temas menos distintos unos de otros con respecto a *BERTopic*, por ende, se restringe el análisis a este último. En general, no se observa una relación significativa con la postura tomada, sin embargo, esto es normal, considerando que se pueden tomar posturas contrapuestas al usar elementos semánticos similares para justificarlas. Para ver ejemplos de justificaciones se incluyen algunas muestras en la Tabla 3. Se observa que el tópico 1 tiene una alta probabilidad de aparición (70,99%) en contraste con los otros. Como indica el ejemplo, la justificación esgrimida en este tópico no muestra necesariamente una complejidad en el desarrollo moral, en ese sentido, destaca la ausencia de principios éticos en las palabras clave para cada tópico. Si bien este modelo no permite un estudio exhaustivo de todos los casos para apoyar la hipótesis acerca de la complejidad en el desarrollo moral, sí conviene tenerla en consideración para futuras investigaciones. Los otros tópicos muestran una baja probabilidad de aparición, pero se pueden notar perspectivas éticas de tipo utilitaristas para escoger la opción de priorizar la salud (porque impactaría a más personas, como muestran los ejemplos).

Tabla 3  
*Caracterización de Tópicos*

<b>Tópico</b>	<b>Palabras clave</b>	<b>Prob</b>	<b>Ejemplo</b>
1	alimento, tradiciones, pueblo	70,99%	Adela debería producir el alimento. Creo que es más importante beneficiar a la gente que posee un déficit vitamínico, ya que es importante resguardar las tradiciones identitarias. Pero eso no está sobre la salud de la población.
2	tradiciones, pueblo, acuerdo	16,8%	Primero sería intentar llegar a un acuerdo con el pueblo, pero si no se puede, si puede salvar personas este fruto entonces es más importante la vida de las personas que una tradición.
3	salud, importante, personas	5,27%	Considero mucho más importante priorizar a la mayor parte de la población por el bien de la salud.
4	proyecto, innovación, si	2,14%	Si bien Adela no tiene limitaciones legales para su proyecto de innovación, sí debería tomar en cuenta la opinión de los pueblos originarios de la zona para tratar de llegar a un acuerdo con ellos, con el fin de respetar las tradiciones y la cultura.
5	Chile, relleno, chilena	1,83%	Ojalá en lo posible se pueda llegar a un acuerdo con los pueblos originarios para respetar aunque sea una parte. Pero definitivamente producir el alimento genera un bien mayor ya que puede ayudar a más del 50% de los niños y adultos mayores en Chile.
6	opinión, mantengo, no	0,97%	Mantengo mi postura

## **Conclusiones**

Respecto a los modelos usados en este trabajo, si bien fueron modelos simples basados en conteos de palabras, la formulación probabilística los dota de la capacidad de clasificar nuevos textos, tanto de manera supervisada como no supervisada. Estas técnicas son además interpretables, característica que es de especial importancia en el contexto del ámbito disciplinar de la ética y su formación. En

el caso específico de *Naive-Bayes*, se tiene tanto una distribución de palabras por clase como una distribución de clase por palabra, lo cual es particularmente útil para analizar el lenguaje empleado en las decisiones de los y las estudiantes.

El desarrollo de otros métodos de aprendizaje profundo capaces de declarar variables (deseablemente, conceptos) que sean destacables a la hora de reflejar la competencia ética, es un trabajo futuro por considerar. En cualquier caso, el presente trabajo muestra una metodología que tiene la capacidad fotografiar la naturaleza de las respuestas y los conceptos empleados para justificarlas. Se espera que esta herramienta ayude a graficar la coherencia de las respuestas cada año, de tal manera que con esta información se puedan tomar decisiones justificadas respecto a la metodología para la enseñanza de la ética en la Facultad de Ciencias Físicas y Matemáticas.

En cuanto a los principales resultados obtenidos a través de los modelos *-Naive-Bayes* y *BERTopic-*, al vincularlos con los principios que subyacen al diferencial estudiado (responsabilidad y respeto), se muestra que, si bien el respeto se tendría en consideración por utilizar la palabra «acuerdo» en sus justificaciones, los y las estudiantes se inclinan hacia razonamientos que priorizan el principio de la responsabilidad (asociado a *producir el alimento contra déficit vitamínico*), en cuanto a la perspectiva de salvar vidas que estén en riesgo en contraste con anteponer un principio de respeto (asociado a *resguardar las tradiciones identitarias*), buscando alternativas a la vitamina vía suplementos o a su obtención a través de la energía del sol. Este hallazgo puede ser representativo del foco de la docencia que está entregando el curso, lo que se relaciona directamente con formar parte de una asignatura de la línea de innovación. Sin embargo, una hipótesis que se plantea es que los y las estudiantes, en términos generales, no utilizan principios éticos subyacentes a su justificación, lo que podría indicar el nivel básico en que se encuentran en el desarrollo del discernimiento ético y, en consecuencia, en la competencia ética.

Este trabajo corresponde a una primera aproximación a las justificaciones de los y las estudiantes, mostrando que se hace necesaria una exploración permanente y diversificada en cuanto a los modelos de procesamiento de lenguaje natural a utilizar.



## Referencias bibliográficas

- Blei, D. M., Ng, A. Y. y Jordan, M. I. (2003). Latent dirichlet allocation. En *Journal of Machine Learning Research*, 3(Jan), 993-1022.
- Cortina, A. (2000). Presentación. El sentido de las profesiones. En A. Cortina y J. Conill, *10 palabras clave en ética de las profesiones*. Ed. Verbo Divino: Navarra.
- Devlin, J., Chang, M.-W., Lee, K. y Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. En *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.
- Garten, J., Boghrati, R., Hoover, J., Johnson, K. M. y Dehghani, M. (2016). Morality Between the Lines: Detecting Moral Sentiment In Text. En *Proceedings of IJCAI 2016 Workshop on Computational Modeling of Attitudes*.
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P. y Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. En *Advances in experimental social psychology (Vol. 47)*. Academic Press. 55, 130.
- Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. En *arXiv preprint arXiv:2203.05794*.
- Kennedy, B., Atari, M., Mostafazadeh Davani, A., Hoover, J., Omrani, A., Graham, J. y
- Dehghani, M. (2021). Moral concerns are differentially observable in language. En *Cognition*, 212, 104696. <https://doi.org/10.1016/j.cognition.2021.104696>
- Ramirez, P., Guerrero, S., Cerda, J., Ross, J., Flores, G. (2022). «La formación ética canalizada mediante la tecnología. Experiencia y resultados preliminares del uso de la herramienta web Ethicapp». *Cuadernos de Beauchef*, 6(2), 73-87.

Xie, J. Y., Hirst, G. y Xu, Y. (2020). Contextualized moral inference. En *arXiv preprint arXiv:2008.10762*. <https://doi.org/10.48550/arXiv.2008.10762>