



## **APLICACIÓN DE MODELOS DE INTELIGENCIA ARTIFICIAL PARA CLASIFICACIÓN Y ANÁLISIS CUALITATIVO DE TEXTOS LIBRES EN ENCUESTAS DIGITALES MASIVAS DE CURSOS DE INGENIERÍA**

Simón Repolt L\*, Universidad de Chile, simondavidandres@gmail.com  
Horacio Buldrini D., Universidad de Chile, hbuldrin@uchile.cl  
Gerardo Baro R., Universidad de Chile, gerardo.baro@ug.uchile.cl  
Ignacio González-Aravena, Universidad de Chile, ignacigonzalez@uchile.cl

### **RESUMEN**

En el presente trabajo se describe un proceso aplicado a respuestas abiertas escritas en texto libre en una encuesta de autopercepción de innovación en cursos masivos de ingeniería, para obtener información que sirva de base para gestionar la docencia de innovación.

El sistema crea de manera comprensiva categorías a partir de palabras claves de conjuntos semánticos generados en función del texto libre de respuestas de encuestas. Por ejemplo, para la pregunta "¿Para qué te es útil la innovación?", una de las categorías generadas es "Creatividad y generación de ideas para el futuro.". Luego, se asocia cada una de las respuestas a una categoría según su contenido. Por ejemplo, la respuesta libre "para poder encontrar nuevas estrategias para solucionar diversos problemas que se presenten..." Se asocia a la categoría "Creatividad y generación de ideas para el futuro." A partir de la preponderancia de cada categoría, se determina la percepción que tienen los estudiantes sobre su experiencia en los cursos, facilitando información para una toma de decisiones efectiva sobre la docencia a través de un sistema automatizado y novedoso que incorpora tecnologías emergentes sobre medios digitales.

**PALABRAS CLAVE:** procesamiento lenguaje natural, LLM, texto libre, innovación, docencia, ingeniería, encuesta

### **1. INTRODUCCIÓN**

En el marco de la automejora institucional continua de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, para poder analizar la progresión en los estudiantes de la competencia de innovación, según el perfil de egreso, se desarrolló una herramienta para medición de la autopercepción de la competencia de innovación en estudiantes de Nivel 1 de la carrera de Ingeniería y Ciencias, que corresponde a primer y segundo año del Plan Común. Particularmente, HÉLICE, el área de formación en ingeniería e innovación, se encarga de tres cursos obligatorios de dicho nivel, los que se desarrollan en modalidad grupal.



La herramienta desarrollada es una iteración del Barómetro de Competencias de Innovación INCODE (ICB) (Penttilä & Kairisto-Mertanen, 2012; Marín-García et al., 2013), una encuesta para la medición de la autopercepción de innovación, parte del proyecto FINCODA (Framework for Innovation Competencies Development and Assessment) de la Unión Europea. Esta se ha validado en múltiples ocasiones y culturas diferentes (Marín-García et al., 2013; Butter & van Beest, 2017; Keinänen et al., 2018; Andreu-Andres et al., 2018; Guzmán-Soria et al., 2022).

El instrumento original se modificó y validó en la Facultad de Ciencias Físicas y Matemáticas, con un modelo de innovación de seis dimensiones: creatividad, pensamiento crítico, iniciativa, colaboración de equipo, moderación de equipo y networking (Baro, Buldrini, & González, 2024). Dicha versión fue revisada por un comité de expertos y pilotada con estudiantes, adaptándose al contexto local y realizándose el despliegue en los cursos del Nivel 1. Se realizaron diversos análisis de las preguntas formuladas en escala Likert, sin embargo, hasta el momento no se había analizado las preguntas de desarrollo.

El propósito de este trabajo es mostrar el sistema desarrollado para automatizar los análisis de las preguntas de texto libre de la encuesta, explorando cómo las herramientas de Inteligencia Artificial, particularmente los modelos de lenguaje de gran tamaño (LLM), pueden facilitar el análisis cualitativo de datos textuales en contextos educativos.

## 2. DESARROLLO

El punto de partida del sistema diseñado son respuestas abiertas en texto libre obtenidas a partir de la encuesta de autopercepción de innovación (por ejemplo: “¿Para qué te es útil la innovación?”). A partir de estas respuestas, el sistema realiza tres funciones principales:

- I. Generación de categorías: El sistema construye categorías comprensibles a partir de las palabras clave que caracterizan los temas emergentes en las respuestas. Por ejemplo, para la pregunta “¿Para qué te es útil la innovación?”, se generan categorías para las respuestas como “Creatividad e innovación transformadora” y “Progreso social mediante ideas y cambios”.
- II. Asignación de respuestas a categorías: Cada respuesta se vincula con la categoría que mejor refleja su contenido semántico. Por ejemplo, la respuesta “para poder encontrar nuevas estrategias para solucionar diversos problemas que se presenten o para satisfacer las necesidades que surgen en el paso del tiempo” se asocia automatizadamente a la categoría “Creatividad e innovación transformadora”.
- III. Estadísticas de ocurrencia: Finalmente, el sistema calcula la frecuencia con que aparece cada categoría dentro de las respuestas, lo que permite obtener una cuantificación de la importancia de cada categoría en las respuestas de texto libre de la encuesta.



Una vez descrito en términos generales lo que el sistema produce, en la siguiente sección se abordará la forma cómo lo hace, es decir, las etapas específicas que permiten transformar texto libre en categorías interpretables.

## 2.1 Diseño metodológico

El sistema implementado procesa respuestas abiertas de encuestas en varias etapas consecutivas, tal como se describe en la figura 1:

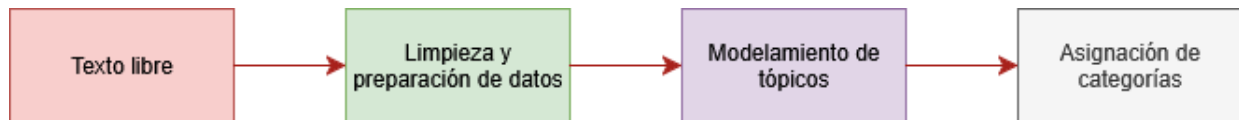


Figura 1: Diagrama general del sistema implementado  
Fuente: Elaboración propia

Se comienza por procesar las respuestas escritas por los estudiantes, almacenadas en una planilla tal como fueron entregadas en la encuesta. Posteriormente, se realiza la limpieza de las respuestas, depurándose para eliminar mayúsculas, caracteres especiales o palabras que no aportan al sentido de la frase. El resultado son versiones simplificadas de cada respuesta, enfocadas en las palabras que realmente llevan significado.

Luego, se implementa el algoritmo de *Alocación Latente de Dirichlet* (LDA) (Blei, 2003), un modelo generativo probabilístico no supervisado diseñado para descubrir patrones ocultos en grandes colecciones de texto, denominados **tópicos** (Blei, 2003). En términos prácticos, este asume que cada respuesta puede pertenecer a varios temas y que cada **tema** se compone de un conjunto característico de palabras. Lo que hace el algoritmo LDA es detectar el listado de palabras que tienden a aparecer juntas en las respuestas y, con ello, descubrir los tópicos que están en el conjunto de datos (respuestas) (Manning, Raghavan, & Schütze, 2008).

Para facilitar la comprensión de sus resultados, se utilizan modelos de lenguaje de gran tamaño (LLM) únicamente en la etapa de denominación de los tópicos, proponiendo nombres interpretables para cada categoría final. De esta forma, el sistema combina la robustez estadística de un modelo no supervisado con la capacidad de interpretación semántica de un LLM, permitiendo transformar respuestas abiertas en categorías comprensibles y útiles para el análisis.



## 2.2 Detalle de la metodología implementada

La figura 2 detalla estos pasos junto a las herramientas y técnicas empleadas en cada etapa, indicando en cursiva el programa utilizado correspondiente.:

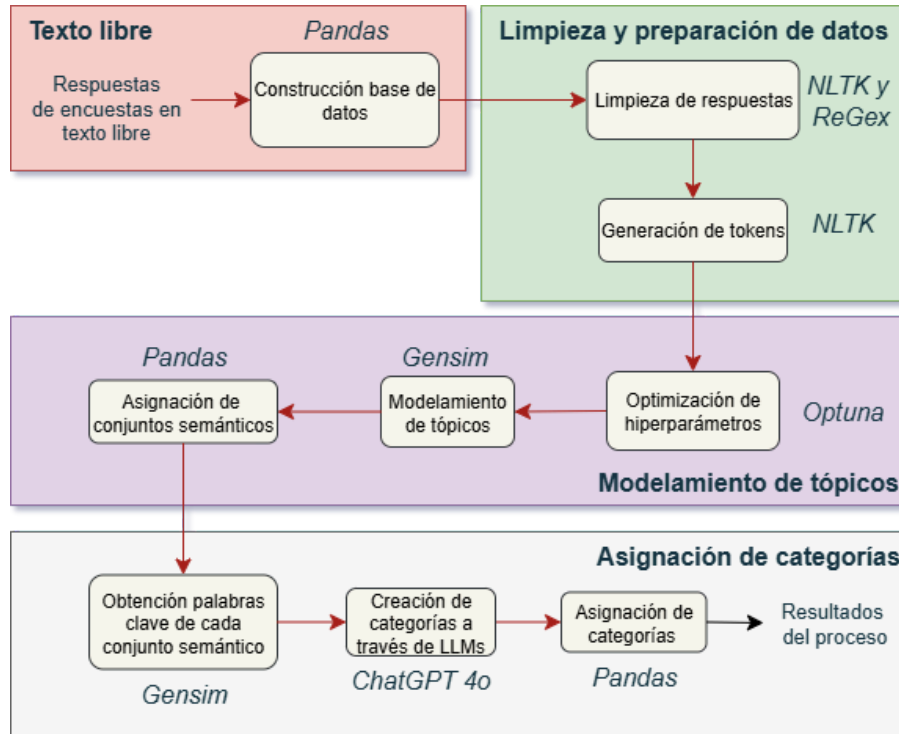


Figura 2: Etapas del procesamiento  
Fuente: Elaboración propia

La Figura 2 presenta en detalle los componentes del sistema, organizados en bloques que representan operaciones específicas: desde la organización inicial de las respuestas, la limpieza lingüística y la preparación de los datos, hasta el modelamiento de tópicos con LDA y la asignación de categorías interpretables mediante un LLM. A continuación, se describe cada bloque de manera individual, explicando su función dentro del flujo de trabajo completo.

### 2.2.1 Construcción de base de datos

Se concatena, mediante la librería *Pandas* (McKinney, 2010), todas las respuestas de las encuestas realizadas y se separa por las preguntas de interés, las cuales en este caso son “¿Para qué te es útil la innovación?” e “Indica para qué te han servido los cursos de innovación.” La base de datos se construye a partir de planillas de Excel que se obtienen directamente de la plataforma donde se realizaron las encuestas.



### 2.2.2 Limpieza de respuestas

Para cada pregunta se eliminan respuestas vacías, y mediante las librerías *Natural Language ToolKit* (NLTK) (Bird, Klein, & Loper 2009) y *re (Regular Expressions)* (Python Software Foundation, 2023) se realiza una limpieza de texto, que sigue las siguientes especificaciones:

- Se eliminan tildes, signos de interrogación, exclamación, puntuación, entre otros.
- Se convierten todas las mayúsculas a minúsculas.
- Se eliminan emoticones y caracteres especiales.
- Se eliminan todas las palabras vacías utilizando una lista de palabras vacías del español proporcionado por la librería NLTK. A este conjunto se le añade otro conjunto personalizado de chilenismos y palabras informales, incluyendo palabras como “cabros”, “goat”, “tá”, “xq”, entre otros (conjunto de elaboración propia).

Por ejemplo, si una respuesta es textualmente “Buena experiENcia en el Ramo de Innovación :) ♥”, tras la limpieza quedaría “buena experiencia ramo innovacion”. Notar que se eliminaron las mayúsculas, símbolos, emoticones, tildes y palabras vacías de la respuesta textual original.

### 2.2.3 Generación de *tokens*

Los *tokens* son las unidades mínimas de análisis que se obtienen al aplicar un proceso que llamamos *tokenización* sobre el texto limpio. En este contexto, la *tokenización* se entiende como el proceso de dividir el texto en fragmentos más simples, usualmente palabras, que posteriormente se representan como **vector** ordenado para el modelamiento (Manning, Raghavan, & Schütze, 2008). Siguiendo el ejemplo anterior, si el texto limpio es "buena experiencia ramo innovacion", entonces el resultado de este proceso se muestra en la Figura 3, donde cada palabra corresponde a un *token*.

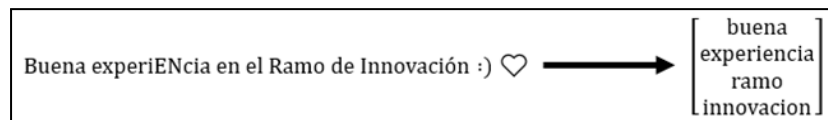


Figura 3: Ejemplo ficticio de conversión de frase en texto libre a vector mediante *tokenización*  
Fuente: Elaboración propia

### 2.2.4 Optimización de hiperparámetros

Un **hiperparámetro** es un valor de configuración de un modelo o sistema que no es inferido o calculado directamente de este, sino que debe definirse antes de ejecutarlo (por ejemplo, el número de tópicos deseados en un modelo LDA). Para que el modelo de tópicos entregue buenos resultados, es necesario encontrar una configuración adecuada de estos hiperparámetros. Para este trabajo se utilizó la librería Optuna (Akiba et al., 2019), que permite



automatizar el proceso de búsqueda de hiperparámetros óptimos. Su funcionamiento puede resumirse en cuatro pasos principales:

- I. Definir el **espacio de búsqueda**: Se establecen los posibles valores de cada hiperparámetro dentro de un rango razonable, entre 4 y 9 tópicos, por ejemplo.
- II. Seleccionar la **métrica de optimización**: En este caso se usó la coherencia de tópicos basada en vectores, una métrica diseñada para aproximar el juicio humano sobre qué tan comprensibles son los grupos de palabras que forman un tópico. Dicha métrica es descrita a continuación:
  - A. Primero, se calcula la información mutua punto a punto normalizada (NPMI) para pares de palabras, a partir de sus co-ocurrencias en el corpus dentro de una ventana de contexto.
  - B. Luego, se mide la similitud entre estos pares utilizando una métrica vectorial (usualmente coseno).
  - C. Finalmente, se promedian los resultados para obtener la coherencia de cada tópico, y la coherencia global del modelo corresponde al promedio de todas ellas. El valor se encuentra entre 0 y 1: un modelo cercano a 0 es mayormente ruidoso, mientras que uno cercano a 1 es muy coherente. La optimización se realiza maximizando esta coherencia (Röder, Both, & Hinneburg, 2015).
- III. Definir las condiciones de búsqueda: Esto incluye cuántas iteraciones de optimización se harán y bajo qué criterio se detendrá la exploración.
- IV. Ejecutar el proceso de optimización: Optuna selecciona hiperparámetros usando el método Tree-structured Parzen Estimator (TPE) (Bergstra et al., 2011). En cada iteración, el algoritmo es:
  - A. Clasifica los resultados anteriores en “buenos” y “malos” (según un umbral, como por ejemplo el 20% superior).
  - B. Aprende dos distribuciones:
    1.  $l(x)$ : Distribución sobre hiperparámetros asociados a buenos resultados.
    2.  $g(x)$ : Distribución sobre hiperparámetros asociados a malos resultados.
  - C. Escoge los siguientes candidatos maximizando la razón  $l(x)/g(x)$ , priorizando combinaciones con más probabilidad de ser “buenas”, pero sin dejar de lado cierta exploración aleatoria.

Al finalizar, Optuna entrega el conjunto de hiperparámetros con mejor desempeño (óptimo o cercano al óptimo) dentro del espacio definido. Con esos valores se procede al entrenamiento del modelo de tópicos. En este proyecto, la optimización se realizó de forma independiente para cada pregunta de la encuesta que se quiso analizar.

### 2.2.5 Modelamiento de tópicos

Para modelar los tópicos en este trabajo se usó LDA. Una vez optimizados sus hiperparámetros, se procede a ejecutar el modelo con dichos valores. Suponiendo que:



- I. Cada **documento** (en este caso, cada respuesta representada como un vector o conjunto de tokens) es una mezcla de varios tópicos.
- II. Cada tópico es una distribución de probabilidad sobre las palabras del vocabulario.

El modelo se basa en una historia generativa:

- I. La distribución de probabilidad asociada a los tópicos se modela a través de una distribución Dirichlet (Bishop, 2006).
- II. Cada documento tiene una distribución de probabilidad sobre los tópicos, también modelada con una distribución Dirichlet, que indica la proporción esperada de cada tópico en ese documento.
- III. Para cada palabra en el documento:
  - A. Se selecciona un tópico de acuerdo con la distribución de tópicos del documento.
  - B. Luego, a partir de ese tópico, se selecciona una palabra siguiendo la distribución de palabras asociada al tópico elegido.

En la práctica, el objetivo no es generar texto, sino inferir a partir de los datos cuáles son esos tópicos y cómo se mezclan en cada documento. Así, el resultado final es:

- I. Una distribución de palabras asociada a cada tópico.
- II. Una distribución de tópicos asociada a cada documento.

En este proyecto, el modelo LDA fue ejecutado con los hiperparámetros óptimos para cada pregunta de la encuesta, obteniéndose la cantidad de tópicos que se detalla más adelante.

### 2.2.6 Asignación de conjuntos semánticos

Una vez ejecutado el modelo LDA para cada pregunta, se obtiene para cada documento una distribución de probabilidad sobre los tópicos. A partir de esta distribución, se selecciona el tópico más probable de cada documento. Luego, este tópico se asocia nuevamente con el texto original (la respuesta de la encuesta) que dio origen al vector de *tokens* utilizado en el modelo. El resultado de esta asociación entre respuesta original, vector de *tokens* y tópico asignado constituye lo que denominamos un **conjunto semántico**.

### 2.2.7 Obtención de palabras clave de cada conjunto semántico

De cada conjunto semántico se extraen 15 palabras más relevantes, determinadas a partir de la distribución de probabilidad de términos generada por el modelo LDA. Estas palabras clave representan el **vocabulario central** de cada tópico.



### 2.2.8 Creación de categorías a través de LLMs

Posteriormente, las palabras clave se utilizan como insumo para un LLM (ChatGPT-4o, en este caso) (OpenAI, 2024). El modelo recibe dichas palabras y sus probabilidades, y propone una etiqueta que describa de forma general el conjunto temático, reduciendo la probabilidad de sobregeneralizar. Para asegurar lo anterior, se emplean las instrucciones descritas en el Anexo 4 al momento de usar el modelo de lenguaje generativo.

Por ejemplo, si para un conjunto identificado como  $id=0$  las palabras más relevantes son *crear*, *creatividad*, *ideas*, *sociedad*, *desarrollar*, *soluciones*, el LLM devuelve una asociación como la que se muestra en la siguiente figura:



Figura 4: Ejemplo de proceso de generación de categorías mediante LLM  
Fuente: Elaboración propia

### 2.2.9 Asignación de categorías

La etapa final del proceso consiste en asignar las categorías (etiquetas) creadas por el LLM a cada respuesta y finalmente exportar estos resultados en una planilla. Por ejemplo, si se considera la respuesta “*El curso fue muy interesante, pero las evaluaciones eran demasiado exigentes y confusas*”, la asociación final que hace el modelo se ve ejemplificada por la siguiente figura:

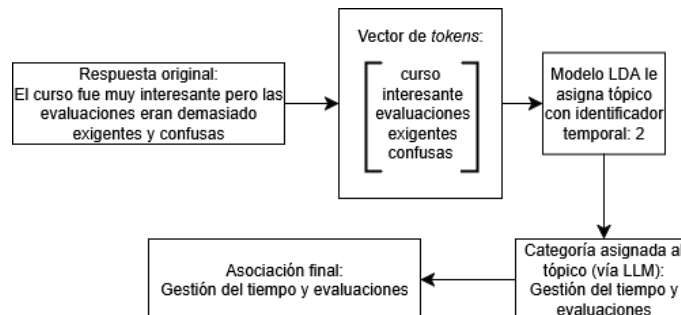


Figura 5: Ejemplo de flujo de procesamiento para una respuesta ficticia  
Fuente: Elaboración propia



### 3. RESULTADOS

El sistema se aplicó sobre una base de datos construida a partir de las respuestas de los estudiantes (Anexo 1.1), cuya cantidad de respuestas válidas se presenta en el Anexo 1.2. Para optimizar los modelos, se efectuó una búsqueda de hiperparámetros con Optuna, realizando 100 iteraciones. Los mejores conjuntos de hiperparámetros tuvieron una coherencia de 63% y 65% respectivamente para cada pregunta analizada, según se detalla en el Anexo 1.3.

Con los hiperparámetros óptimos, se ejecutaron los modelos LDA obteniendo las distribuciones de probabilidad para cada tópico. En particular, se establecieron 11 tópicos para la primera pregunta y 12 tópicos para la segunda. Los conjuntos de hiperparámetros, junto con sus explicaciones y coherencia, se incluyen en el Anexo 2. Las palabras clave de cada conjunto semántico fueron posteriormente asociadas a categorías generadas mediante un LLM, como se puede ver en la tabla 1, mostradas según orden de generación:

Tabla 1: Categorías generadas para cada pregunta

Pregunta	Categorías	Frec.
<b>¿Para qué te es útil la innovación?</b>	Creatividad y generación de ideas para el futuro.	14.31%
	<b>Innovación para mejorar la sociedad y la vida cotidiana.</b>	<b>82.69%</b>
	Propuestas efectivas con visión crítica y comunitaria.	0.73%
	Trabajo en equipo y optimización de procesos sociales.	2.27%
<b>Indica para qué te han servido los cursos de innovación.</b>	Ampliación de perspectivas y pensamiento flexible.	0.14%
	Aprendizaje a través del curso y experiencias de clase.	1.84%
	<b>Organización personal y proyección a futuro.</b>	<b>42.34%</b>
	Pensamiento crítico y aplicación de métodos innovadores.	1.32%
	Resolución de problemas mediante nuevas formas.	0.07%
	<b>Trabajo colaborativo, creatividad y resolución de problemas.</b>	<b>54.29%</b>

Fuente: Elaboración propia

En la primera pregunta, la gran mayoría de respuestas (82.69%) se concentró en la categoría “Innovación para mejorar la sociedad y la vida cotidiana”. Esto muestra que los estudiantes conciben la innovación principalmente como un medio de impacto social, más que como una herramienta individual.



**XXXVII CONGRESO CHILENO DE EDUCACIÓN EN INGENIERÍA 2025**  
PROYECCIÓN DE LAS TECNOLOGÍAS DIGITALES EN LA FORMACIÓN EN INGENIERÍA:  
LA EDUCACIÓN EN MODALIDAD PRESENCIAL, HÍBRIDA Y VIRTUAL  
Concepción, 8 al 10 de octubre 2025

En la segunda pregunta, la distribución fue más equilibrada. La categoría principal, “Trabajo colaborativo, creatividad y resolución de problemas” (54.29%), convive con otras que destacan principalmente el pensamiento crítico, la organización personal y la ampliación de perspectivas. Esto refleja que los cursos son valorados por fomentar habilidades prácticas y diversas. En conjunto, los resultados sugieren una diferencia clara: mientras la innovación en abstracto se asocia a un ideal colectivo y social, los cursos de innovación son percibidos como instancias para el desarrollo de habilidades aplicadas y multifacéticas.

## CONCLUSIONES

La herramienta de análisis de respuestas que se ha presentado en este trabajo, ha permitido extraer información útil. A partir de respuestas en texto libre a preguntas de desarrollo se ha podido obtener información significativa de la percepción de los estudiantes respecto de algunas de sus habilidades y competencias personales asociadas con estos cursos.

Específicamente, los resultados que se han obtenido hasta el momento permiten determinar que los estudiantes tienen una autopercepción muy fortalecida de la dimensión de trabajo en equipo, indicando que estos cursos están contribuyendo de manera importante a lograr la instalación o consolidación en los estudiantes de esa dimensión específica de la innovación. Como las menciones a elementos de las demás dimensiones de la innovación han sido menos frecuentes que la de trabajo en equipo, en base a estos resultados se podría recomendar reforzar esas otras dimensiones en los contenidos de los cursos.

Por otra parte se podría extender este mecanismo de consulta espontánea mediante otras preguntas con respuestas de texto libre buscando que puedan servir para identificar los aspectos que conviene reforzar en los contenidos o metodologías de los cursos.

El complemento a la encuesta que se ha desarrollado en este trabajo es una herramienta que fortalece el análisis ya existente de las encuestas realizadas por medios digitales en cursos masivos permitiendo extraer información importante de las respuestas en texto libre. Ello abre la posibilidad de contar con una alternativa a los métodos tradicionales de preguntas con respuestas preconfiguradas (alternativas, Likert, etc.).

Creemos que contar con éste análisis de la percepción estudiantil con ayuda de Inteligencia Artificial en varias instancias del desarrollo estudiantil, permitirá institucionalizar un ciclo de mejora continua de estos cursos de innovación fortaleciendo los procesos de aseguramiento de calidad: recolectar datos, analizarlos con rigor, retroalimentar a docentes y rediseñar el currículo cuando la evidencia recogida lo haga aconsejable. Eso facilitará el alineamiento de los cursos de innovación con los desafíos de la educación de ingeniería.



## ANEXOS

### Anexo 1: Tablas

#### Anexo 1.1

Tabla 2: Cantidad de respuestas por instancia de la encuesta

	Instancia 1	Instancia 2	Instancia 3	Instancia 4	Total
<b>Respuestas</b>	284	722	760	65	1.831

Fuente: Elaboración propia

#### Anexo 1.2

Tabla 3: Cantidad de respuestas limpias por instancia y por pregunta de interés de la encuesta

Pregunta	Instancia 1	Instancia 2	Instancia 3	Instancia 4	Total
"¿Para qué te es útil la innovación?"	182	522	615	43	1362
"Indica para qué te han servido los cursos de innovación."	185	524	605	44	1358

Fuente: Elaboración propia

#### Anexo 1.3

Tabla 4: Coherencia obtenida para cada pregunta e iteración donde se alcanzó

Pregunta	Coherencia basada en vectores	Iteración donde se alcanzó
¿Para qué te es útil la innovación?	0.634	94
Indica para qué te han servido los cursos de innovación.	0.645	67

Fuente: Elaboración propia



## Anexo 2: Hiperparámetros óptimos por pregunta.

A continuación se definen los hiperparámetros del modelo LDA a optimizar a través de Optuna:

1. *No\_below* :Número mínimo de documentos en los que un token debe aparecer para ser incluido en el diccionario. Filtra palabras demasiado raras, reduciendo el ruido presente en el vocabulario.
2. *No\_above*: Umbral máximo de frecuencia relativa de aparición de un token en el corpus (proporción de documentos). Filtra palabras demasiado frecuentes, como stopwords o términos no discriminatorios.
3. *Num\_topics*: Número de tópicos latentes a inferir. Valores bajos generan tópicos más amplios y generales; valores altos producen tópicos más específicos, aunque con mayor riesgo de solapamiento.
4. *Alpha*: Parámetro Dirichlet que controla la distribución de tópicos por documento.
  - Valores bajos: documentos dominados por pocos tópicos.
  - Valores altos: documentos con mezcla de varios tópicos.  
En Gensim puede configurarse como "symmetric", "asymmetric" o "auto".
5. *Eta*: Parámetro Dirichlet que controla la distribución de palabras por tópico.
  - Valores bajos: cada tópico se concentra en pocas palabras dominantes.
  - Valores altos: los tópicos son más uniformes y difusos. Puede definirse como "symmetric" o "auto".
6. *Passes*: Número de pasadas completas sobre el corpus durante el entrenamiento. Un mayor número de pasadas incrementa la estabilidad de los resultados, aunque a mayor costo computacional.
7. *Iterations*: Número de iteraciones internas por lote durante la inferencia. A mayor número de iteraciones, mejor ajuste de los parámetros latentes, pero con mayor tiempo de cómputo.
8. *Chunksize*: Número de documentos procesados por lote en cada actualización del modelo.
  - Valores grandes: estimaciones más estables y menos ruido.
  - Valores pequeños: más variabilidad, pero mayor frecuencia de actualización.
9. *Minimum\_probability*: Probabilidad mínima para reportar un tópico en la distribución de un documento. Filtra tópicos con contribución marginal, generando salidas más concisas.

En la tabla 5 se especifican los valores óptimos obtenidos a través de Optuna para cada pregunta:



Tabla 5: Valores óptimos por hiperparámetro de Optuna

Hiperparámetro	¿Para qué te es útil la innovación?	Indica para qué te han servido los cursos de innovación.
<i>No_below</i>	3	5
<i>No_above</i>	0.1	0.3
<i>Num_topics</i>	11	12
<i>Alpha</i>	<i>asymmetric</i>	<i>asymmetric</i>
<i>Eta</i>	<i>auto</i>	0.013
<i>Passes</i>	125	140
<i>Iterations</i>	450	2400
<i>Chunksize</i>	200	1150
<i>Minimum_probability</i>	0.189	0.068

Fuente: Elaboración propia

### Anexo 3: Especificaciones computacionales

El sistema en su completitud fue ejecutado en un computador de escritorio con Windows 11 pro, utilizando un procesador Intel I7 - 6700K operando a 4 GHz y con 16 GB de RAM DDR4 operando a 2.4 GHz. Se utilizó la tarjeta gráfica integrada del procesador para los cómputos.



#### Anexo 4: Instrucciones para ChatGPT 4o

A continuación se presentan las instrucciones dadas a ChatGPT para generar los nombres de las categorías:

“Tengo la salida de un modelo LDA o similar, donde para cada pregunta tengo varios tópicos, y cada tópico está representado por una lista de palabras clave con su probabilidad.

Quiero que generes nombres descriptivos en español para cada tópico, en formato de diccionario ({número\_tópico: "nombre descriptivo"}), basándote en las palabras más representativas.

Si varios tópicos tienen exactamente las mismas palabras clave o son muy similares (por ejemplo, ruido), usa el mismo nombre para todos ellos.

No simplifiques demasiado si hay matices distintos; en ese caso, crea nombres diferenciados.

Dame los nombres en español. Si quieres, puedes sugerir también versiones resumidas para dashboards, y si te lo pido, puedes darme el resultado como .json, .py o .csv.”



## REFERENCIAS

Akiba T., Sano S., Yanase T., Ohta T., & Koyama M.. (2019). *Optuna: A Next-generation Hyperparameter Optimization Framework*. En Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19) (pp. 2623–2631). Association for Computing Machinery. <https://doi.org/10.1145/3292500.3330701>

Andreu-Andres, M. Á., Gonzalez-Ladron-de-Guevara, F. R., Garcia-Carbonell, A., & Watts-Hooge, F. (2018). *Contrasting innovation competence FINCODA model in software engineering: Narrative review*. Journal of Industrial Engineering and Management, 11(4), 715-734. <https://doi.org/10.3926/jiem.2656>

Baro, G., Buldrini, H., González, I. (2024). *Diseño de herramienta para medición de la competencia de innovación en estudiantes de ingeniería*. En XXXVI Congreso Chileno de Educación en Ingeniería 2024.

Bergstra, J., Bardenet, R., Bengio, Y., y Kégl, B. (2011). *Algorithms for hyper-parameter optimization*. En Proceedings of the 24th International Conference on Neural Information Processing Systems (NIPS'11) (pp. 2546–2554). Curran Associates Inc. <https://proceedings.neurips.cc/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf>

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. O'Reilly Media, Inc.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). *Latent Dirichlet Allocation*. Journal of Machine Learning Research, 3, 993–1022.

Butter, R., & van Beest, W. (2017). *Validating the FINCODA Innovation Barometer. Psychometric validation of a tool for innovation competencies development and assessment*. en FINCODA Project Conference Proceedings.

<https://www.fincoda.eu/fincoda-blog/2017/3/27/psychometric-validation-of-a-tool-for-innovation-competencies-development>



**XXXVII CONGRESO CHILENO DE EDUCACIÓN EN INGENIERÍA 2025**  
PROYECCIÓN DE LAS TECNOLOGÍAS DIGITALES EN LA FORMACIÓN EN INGENIERÍA:  
LA EDUCACIÓN EN MODALIDAD PRESENCIAL, HÍBRIDA Y VIRTUAL  
Concepción, 8 al 10 de octubre 2025

Guzmán-Soria, E., De la Garza-Carranza, M. T., Herrera-Laguna, V. Y., & Sánchez-Mojica, K. Y. (2022). *El autoliderazgo, la autoeficacia emprendedora y las competencias innovadoras de la generación centenal en México*. Investigación y Ciencia de la Universidad Autónoma de Aguascalientes, 30. <https://doi.org/10.33064/iycuaa2022874080>

Keinänen, M., Ursin, J., & Nissinen, K. (2018). *How to measure students' innovation competences in higher education: Evaluation of an assessment tool in authentic learning environments*. Studies in Educational Evaluation, 58, 30-36.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Marín-García, J. A., Pérez-Peñalver, M. J., & Watts, F. (2013). *How to assess innovation competence in services: The case of university students*. Dirección y Organización, 50, 48-62. <https://revistadyo.es/DyO/index.php/dyo/article/view/431/451>

McKinney, W. (2010). *Data structures for statistical computing in Python*. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 51–56).

OpenAI. (2024). *ChatGPT-4o* (Version May 2024) [Large language model]. OpenAI. <https://chat.openai.com>

Penttilä, T., & Kairisto-Mertanen, L. (2012). *Innovation Competence Barometer ICB – A Tool For Assessing Students' Innovation Competences As Learning Outcomes In Higher Education*. INTED2012 Proceedings, 6347-6351. <https://library.iated.org/view/PENTTILA2012INN>

Python Software Foundation. (2023). *re — Regular expression operations*. En Python 3.11.5 documentation. <https://docs.python.org/3/library/re.html>

Röder, M., Both, A., y Hinneburg, A. (2015). *Exploring the Space of Topic Coherence Measures*. En Actas de la octava conferencia internacional ACM sobre búsqueda y minería en la web (pp. 399-408). Association for Computing Machinery. <https://doi.org/10.1145/2684822.2685324>