



Universidad del Desarrollo
Facultad de Ingeniería

ANÁLISIS TEMÁTICO Y DE SENTIMIENTO DEL DISCURSO PARLAMENTARIO
CHILENO

Un enfoque de procesamiento del lenguaje natural aplicado a sesiones del Senado
(2010-2025)

POR: VICENTE IGNACIO GALLARDO HENRIQUEZ

Proyecto de grado presentado a la Facultad de Ingeniería de la Universidad del
Desarrollo para optar al grado académico de Magíster en Data Science

PROFESOR GUÍA:

Dra. María Paz Raveau

[Diciembre, 2025]

[SANTIAGO]

A mi esposa y a mi familia, por ser mi principal sostén durante este proceso académico.

AGRADECIMIENTO

Agradezco a la Dra. María Paz Raveau por su orientación y observaciones a lo largo de desarrollo de este trabajo, las cuales contribuyeron a mejorar su enfoque y claridad metodológica.

Asimismo, agradezco a mi esposa por su apoyo constante, comprensión y acompañamiento durante este proceso, especialmente en las etapas de mayor dedicación y exigencia.

ÍNDICE

RESUMEN.....	1
1. INTRODUCCIÓN	2
2. TRABAJO RELACIONADO	4
3. HIPÓTESIS Y OBJETIVOS.....	5
4. DATOS Y METODOLOGÍA.....	6
4.1. DATOS	6
4.2. METODOLOGÍA.....	7
4.2.1 <i>Detección de unidades léxicas compuestas</i>	8
4.2.2 <i>Modelo Word2Vec entrenado sobre el corpus</i>	9
4.2.3 <i>Modelo FastText preentrenado cc.es.300</i>	9
4.2.4 <i>Expansión semántica de los diccionarios</i>	10
4.2.5 <i>Clasificación temática del corpus</i>	10
4.2.6 <i>Análisis de sentimiento basado en modelos Transformers</i>	12
5. RESULTADOS	14
5.1 <i>Desempeño inicial del clasificador basado en diccionarios</i>	15
5.2 <i>Diagnóstico y refinamiento de errores</i>	16
5.3 <i>Distribución temática del discurso parlamentario</i>	20
5.4 <i>Validación y selección del modelo de análisis de sentimiento</i>	23
5.5 <i>Análisis del sentimiento del discurso parlamentario</i>	27
6. CONCLUSIONES	36
BIBLIOGRAFÍA	39

Resumen

Este proyecto analiza el discurso parlamentario chileno desde una doble dimensión temática y emocional, aplicando técnicas de procesamiento de lenguaje natural a un corpus de 1.674 sesiones del Senado chileno entre 2010 y 2025, con más de 41 millones de palabras. Para la clasificación temática se construyó un diccionario maestro en cinco ejes de política pública (economía, educación, salud, medioambiente y seguridad), iniciado desde términos semilla extraídos de literatura especializada y expandido semánticamente mediante embeddings (Word2Vec entrenado sobre el corpus y FastText preentrenados), seguido por un proceso iterativo de depuración y validación manual para reducir activaciones falsas y mejorar especificidad temática.

En resultados, la clasificación aplicada al corpus completo muestra una agenda temática relativamente equilibrada con mayor presencia de economía (22,52%) y seguridad (21,33%), seguidas por salud (19,18%), educación (17,44%) y medioambiente (9,02%), mientras que “Otros” representa 10,51% de sesiones sin predominancia temática según los criterios definidos.

Aplicando a las sesiones con tópico asignado (excluyendo “Otros”), el análisis de sentimiento sobre 1.498 sesiones segmentadas en 140.977 bloques muestra predominio de neutralidad (~61%), con una proporción relevante de negatividad (~32%) y un componente positivo minoritario (~7%), patrón coherente con un registro legislativo técnico y argumentativo. Finalmente, la evolución temporal indica variaciones sistemáticas del tono emocional, destacando un mínimo en 2020 y una moderación posterior, sugiriendo sensibilidad del discurso a coyunturas político-institucionales.

En conjunto, el estudio confirma que el discurso parlamentario chileno presenta estructuración temática identificable mediante enfoques léxicos validados y variación emocional sistemática medible con modelos de lenguaje adecuados al dominio, aportando una metodología reproducible para caracterizar prioridades temáticas y clima discursivo a gran escala.

1. Introducción

El estudio del discurso político constituye una herramienta fundamental para comprender cómo se configuran las relaciones de poder, la legitimidad institucional y las dinámicas ideológicas en el ámbito público. En el caso de Chile, el debate parlamentario ofrece un registro para observar estos procesos, por ejemplo, en las sesiones del Senado se documentan de forma sistemáticas las discusiones, desacuerdos y acuerdos que moldean la política nacional.

Tradicionalmente, la investigación sobre el discurso parlamentario ha estado dominada por enfoques cualitativos centrados en el análisis pragmático o retórico (Núñez, 2012), lo que ha permitido identificar estrategias lingüísticas y de interacción relevantes, pero con una cobertura limitada frente a la gran cantidad de material disponible. La incorporación de métodos computacionales y herramientas de Procesamiento de Lenguaje Natural (PLN) ha ampliado las posibilidades de análisis, permitiendo examinar de manera sistemática patrones léxicos ideológicos y emocionales en grandes corpus textuales (Horn y Balenzuela, 2025).

En este contexto, el presente trabajo propone un análisis del discurso parlamentario chileno a partir del corpus de sesiones del Senado, con el objetivo de estudiar tanto el contenido temático como el tono emocional de los debates. En particular, se busca identificar la presencia y evolución de cinco tópicos centrales—economía, educación, salud, medioambiente y seguridad- y examinar cómo estos se relacionan con la carga afectiva del discurso político.

Para ello, se construyeron diccionarios temáticos a partir de literatura científica, creando listas de términos semillas, los cuales se expanden automáticamente con un modelo Word2Vec entrenado sobre el propio corpus parlamentario y el Modelo FastText preentrenado cc.es.300 en español. Posteriormente, ambos conjuntos de embeddings se integran mediante ponderaciones y se sometieron a un proceso de depuración manual con el fin de maximizar la cobertura y precisión de la clasificación temática.

Sobre esta base, se evaluaron modelos preentrenados de análisis de sentimiento basado en Transformers, específicamente un modelo fine-tuning de BETO y RoBERTuito Sentiment, modelo de la librería Pysentimiento, con el fin de determinar cuál ofrece un mejor desempeño en el dominio legislativo. A partir de una muestra, se comparó su rendimiento y se seleccionó el modelo BETO como clasificador principal. Con este modelo se estimó la orientación emocional del discurso y se analiza su variación según tópico y periodo legislativo, combinando medidas agregadas (mediana y promedio del

sentimiento, distribución porcentual) con indicadores de evolución intra-sesión (polaridad por deciles de cada debate).

2. Trabajo Relacionado

El análisis del discurso parlamentario ha sido abordado tradicionalmente desde la lingüística y la comunicación política, con especial énfasis en los aspectos retóricos, pragmáticos y estratégicos de la interacción legislativa. En el contexto de Chile, Núñez (2012) realizó un estudio en esta línea, examinando las estrategias de descortesía en el discurso parlamentario y demostrando que la confrontación verbal y la cortesía negativa cumplen un rol político relevante al reforzar identidades partidarias y posicionamiento ideológicos. Este enfoque ha permitido caracterizar la dinámica del debate legislativo como un espacio híbrido entre formalidad institucional y conflicto discursivo.

Posteriormente, en otros trabajos como De Cock (2012), se ha ampliado el análisis hacia la comparación entre registros discursivos, evidenciando un proceso de mediatización del lenguaje político, donde el discurso legislativo adopta rasgos del habla cotidiana y de los medios de comunicación, tales como la personalización del discurso, el uso de expresividad emocional y una orientación persuasiva dirigida al electorado.

Además, la incorporación de Procesamiento de Lenguaje natural ha permitido automatizar tareas que antes requerían análisis manual intensivo. En esta línea, Horn y Balenzuela (2025) presentan una metodología para el análisis automatizado de discursos políticos en América latina mediante técnicas como la detección de frecuencia léxica, análisis de sentimiento y agrupamiento temático. Paralelamente, en el campo del modelamiento lingüístico en español, el desarrollo de BETO (Cañete, 2020) marcó un hito al proveer uno de los primeros modelos de Transformer preentrenado exclusivamente en corpus en español. Más recientemente, la librería Pysentimiento (Pérez, 2021;2023) introdujo modelos optimizados para la clasificación afectiva, como RoBERTuito Sentiment, ampliamente validados en tareas de análisis emocional.

Estos avances metodológicos han consolidado un marco técnico que facilita el estudio computacional del discurso parlamentario en español. El presente trabajo se trabaja en esta línea, integrando diccionarios temáticos construidos y expandidos mediante modelos semánticos en español junto con clasificadores de sentimiento basados en Transformers.

3. Hipótesis y Objetivos

El presente proyecto parte de la premisa que el discurso parlamentario chileno no solo varía en función del contenido temático, sino que también en su carga emocional, siendo ambos elementos indicadores de prioridades políticas, estilos comunicativos y dinámicas institucionales. Dado que el análisis manual del corpus resulta inviable por su magnitud, se propuso explorar estos patrones mediante técnicas de procesamiento de lenguaje natural integradas a un enfoque léxico y de modelos de representación contextual.

Bajo este marco, se plantea que la estructuración temática del discurso puede ser identificada con un nivel razonable de precisión mediante un diccionario especializado, construido a partir de semillas conceptuales, expandido con técnicas de similaridad semántica y finalmente una revisión manual. Este enfoque permitirá detectar variaciones en la agenda discursiva y facilitar el análisis comparado entre periodos legislativos.

En este contexto, los objetivos específicos del proyecto son:

1. Construir un diccionario temático ampliado para los ejes del debate parlamentario (economía, educación, salud, medioambiente y seguridad), combinando un conjunto de términos base con técnicas de expansión automática mediante modelos de embeddings.
2. Clasificar el corpus de intervenciones parlamentarias por tópico utilizando dicho diccionario, evaluando precisión (accuracy) y cobertura como métricas cuantitativas principales.
3. Aplicar modelos de análisis de sentimiento para estimar la orientación emocional del discurso por tema, probando inicialmente distintas alternativas y seleccionando aquella que entregue un desempeño más adecuado para el dominio legislativo.
4. Examinar cómo varía el uso de emoción y la distribución temática en el tiempo entre los distintos tópicos, identificando patrones que contribuyan a caracterizar el estilo discursivo del Congreso en Chile.

4. Datos y Metodología

4.1. Datos

El corpus utilizado corresponde a 1.674 sesiones del senado de Chile, desde el año 2010 hasta septiembre de 2025, obtenidas desde los registros públicos disponibles (<https://opendata.camara.cl>). Este periodo fue seleccionado porque abarca dos ciclos completos de alternancia política en Chile, lo que permite comparar el discurso parlamentario bajo administraciones de distinto signo y evita sesgos a un periodo político único. Asimismo, el intervalo temporal ofrece un horizonte temporal suficientemente largo para detectar tendencias estables en el lenguaje legislativo.

Cada sesión contiene el texto íntegro de las intervenciones, además de metadatos como:

- Fecha de realización
- Número de sesión

Tras un proceso de consolidación y estandarización, todas las sesiones fueron integradas en un único archivo en formato parquet, incorporando tanto los metadatos como las transformaciones textuales preliminares aplicadas durante el proyecto.

El corpus completo supera los 41 millones de palabras, lo que convierte en un insumo de suficiente escala para la aplicación de técnicas de procesamiento de lenguaje natural basada en modelos distribucionales y de lenguaje avanzado, permitiendo realizar análisis temáticos y afectivos con suficiente robustez estadística.

Métrica	Valor
Número total de sesiones	1.674
Palabras totales del corpus	41.137.173
Palabras promedio por sesión	24.574
Sesión más corta (palabras)	45
Sesión más larga (palabras)	140.956
Sesión más antigua	11-03-2010
Sesión más reciente	16-09-2025

Tabla N°1: Tabla descriptiva del corpus

La tabla N°1 presenta las principales estadísticas descriptivas del corpus legislativo analizado, incluyendo su tamaño total, la distribución de palabras por sesión y rango temporal cubierto. Estas métricas permiten dimensionar la escala del conjunto de datos y respaldan la idoneidad del corpus para la aplicación de técnicas de análisis del discurso.

4.2. Metodología

Con el propósito de preparar el texto para su análisis mediante modelos de representación semántica y técnicas de clasificación, se desarrolló un proceso de preprocesamiento cuidadosamente diseñado para mantener la correcta lingüística del discurso parlamentario. En primer lugar, se aplicó una normalización básica que incluyó pública la conversión a minúsculas y la eliminación de saltos de línea, respetando las tildes y sin eliminar stopwords ni signos de puntuación relevantes, dado que estos aportan información gramatical en textos legislativos.

Luego, se realizó la tokenización y lematización del corpus utilizando la librería Stanza con el modelo preentrenado para español aplicada mediante procesamiento por lotes de 200 documentos. Este procedimiento permitió transformar cada intervención parlamentaria en una secuencia de lemas, estandarizando las variantes observadas en el discurso legislativo (por ejemplo, ‘estructuras críticas’ por ‘estructura crítico’).

La identificación temática se apoyó en la creación de un diccionario maestro compuesto por cinco áreas de política pública: economía, educación, salud, medioambiente y seguridad. Cada diccionario comenzó con un conjunto de términos semilla definidos a partir de literatura especializada. Para cada área temática se consultaron aproximadamente 20 artículos académicos de bases de datos como SciELO y Google Scholar, utilizando combinaciones de búsqueda para recuperar vocabulario conceptual y terminología de política pública. Las consultas se estructuraron mediante patrones del tipo:

- Seguridad + Narcotráfico
- Educación + Política Educativa
- Medioambiente + Cambio climático
- Economía + Crecimiento económico
- Salud + Salud pública

Estas búsquedas permitieron identificar conceptos, asegurando que los términos semillas representaran adecuadamente el discurso técnico y académico asociado a cada tópico. Posteriormente cada diccionario fue expandido semánticamente mediante dos modelos complementarios. La combinación permitió equilibrar precisión contextual y amplitud léxica. Tras la expansión automática y la asignación de puntuaciones de similitud, se aplicó una revisión manual para eliminar términos excesivamente genéricos o ambiguos (por ejemplo, control, frontera, estrategia). El resultado final fue un diccionario maestro con aproximadamente 400 términos distribuidos entre cinco tópicos, incluyendo lemas simples y (n-gram), que capturan adecuadamente la terminología política asociada a cada tópico.

La metodología aplicada integra técnicas clásicas de procesamiento de lenguaje natural, modelos distribucionales entrenados sobre el corpus legislativo y modelos de lenguaje basados en Transformers para análisis de sentimiento. A continuación, se describen los principales componentes del procedimiento analítico.

4.2.1 Detección de unidades léxicas compuestas

Con el fin de capturar adecuadamente expresiones recurrentes y relevantes en el discurso parlamentario, se aplicó un proceso de detección de bigramas y trigramas sobre el corpus completo de sesiones del Senado, utilizando la herramienta Gensim Phrases. Este método permitió identificar combinaciones de palabras que coocurren con alta frecuencia y que representan conceptos con significado dentro del dominio parlamentario, tales como ‘crimen_organizado’, ‘salud_mental’ o ‘seguridad_publica’.

La detección de estas expresiones se realizó sobre los textos previamente normalizados y lematizados, lo que permitió reducir variaciones superficiales y favorecer la identificación estable de combinaciones léxicas relevantes. Para ello, se emplearon los siguientes parámetros:

- min_count = 5
threshold = 10

Buscando equilibrar la detección de expresiones relevantes con la exclusión de combinaciones no representativas o poco informativas.

Es importante señalar que este procedimiento no reemplaza los términos simples definidos en el diccionario semilla, sino que los complementa. El diccionario temático inicial, construido a partir de términos definidos manualmente para cada eje temático, se ve enriquecido mediante la incorporación de términos compuestos (bigramas y trigramas)

detectados automáticamente en el corpus parlamentario. De este modo, el vocabulario final integra tantos términos simples (unigramas) como expresiones compuestas características del discurso legislativo, lo que permite una representación más adecuada del discurso y contribuye a mejorar tanto la posterior expansión semántica mediante modelos de embeddings como la precisión de la clasificación temática.

4.2.2 Modelo Word2Vec entrenado sobre el corpus

Con el propósito de capturar relaciones semánticas específicas del discurso parlamentario, se entrenó un modelo Word2Vec utilizando el corpus completo de sesiones del Senado, previamente normalizado y lematizado, e incorporando las unidades léxicas compuestas (bigramas y trigramas) detectadas en la etapa anterior. De este modo, el vocabulario incluyó tanto términos simples (unigramas) como expresiones compuestas propias del registro legislativo.

El modelo fue entrenado exclusivamente sobre el corpus parlamentario, con el objetivo de capturar relaciones semánticas acorde al uso lingüístico y conceptual propio del debate legislativo. Esta decisión permite que las representaciones aprendidas reflejen de manera más precisa el vocabulario técnico, político e institucional característico de las sesiones del Senado, en lugar de depender de asociaciones propias de corpus de carácter general.

Se utilizó la arquitectura skip-gram debido a su capacidad para aprender relaciones semánticas finas, con los siguientes parámetros:

- vector_size = 300
- sg = 1
- window = 5
- min_count = 5
- epochs = 15

Este modelo permitió estimar similitudes ajustadas al uso lingüístico propio del Senado, especialmente relevante para detectar términos técnicos, políticos o institucionales que no suelen aparecer en grandes corpus generalistas.

4.2.3 Modelo FastText preentrenado cc.es.300

Como complemento al modelo Word2Vec entrenado localmente, se incorporó el modelo FastText cc.es.300 preentrenado por Facebook sobre más de dos mil millones de palabras. Este modelo incorpora representaciones sub-léxicas mediante character n-grams lo que permite:

- reconocer palabras derivadas
- cubrir términos no presentes en el corpus legislativo
- incorporar semántica generalista del español

FastText compensa la menor cobertura del Word2Vec entrenado localmente, especialmente en casos donde aparecen términos poco frecuentes, tecnicismos legislativos aislados o variaciones no lematizadas totalmente.

4.2.4 Expansión semántica de los diccionarios

Cada termino semilla fue expandido mediante ambos modelos distribucionales. Para cada palabra o bigrama se obtuvieron

- similitud utilizando el modelo de Word2Vec entrenado sobre el corpus
- similitud utilizando el modelo de FastText preentrenado
- versiones normalizadas de ambas medidas de similitud

Posteriormente se generó un puntaje combinado, definido como:

$$Puntaje = 0,5 \cdot similitud_{W2V} + 0,5 \cdot similitud_{FT}$$

Este esquema de ponderación fue elegido tras evaluar coberturas y ejemplos. Word2Vec aporta precisión contextual, mientras que FastText aporta cobertura. La combinación permite equilibrar ambos efectos.

Por último, se aplicó una revisión manual para eliminar falsos positivos semánticos y términos excesivamente generales o ambiguos, asegurando que el diccionario final reflejara adecuadamente la terminología específica de cada tópico.

4.2.5 Clasificación temática del corpus

Antes de aplicar la clasificación temática a la totalidad del corpus, se realizó una validación preliminar mediante una muestra aleatoria de 350 sesiones parlamentarias. Esta muestra fue inicialmente clasificada utilizando el diccionario maestro, asignando a cada sesión un tópico dominante en función de las coincidencias léxicas.

A partir de esta primera clasificación, se seleccionó una submuestra de 52 sesiones para una revisión manual detallada, compuesta por 10 sesiones por cada uno de los cinco tópicos temáticos (economía, educación, salud, medioambiente y seguridad), además de 2 sesiones clasificadas inicialmente como “otros”. Esta selección permitió evaluar de

manera equilibrada el comportamiento del diccionario en cada tópico y examinar explícitamente los casos sin asignación temática dominante.

La revisión manual de esta submuestra consistió en contrastar el tópico asignado automáticamente con la lectura del contenido de cada sesión, evaluando si la clasificación reflejaba adecuadamente el foco principal del debate. Este proceso permitió identificar errores sistemáticos de clasificación, asociados principalmente a la presencia de términos excesivamente generales o transversales que activaban tópicos dominantes sin que estos representaran el eje central de la discusión.

Como resultado de esta revisión manual focalizada, se eliminaron o ajustaron múltiples términos ambiguos y se refinaron los criterios de clasificación, reduciendo la incidencia de falsos positivos y mejorando la precisión del diccionario antes de su aplicación definitiva. Sobre la base de la clasificación automática inicial de la muestra y de la revisión manual de la submuestra, se definieron umbrales mínimos específicos por tópico, con el fin de evitar que la presencia de uno o dos términos aislados fuera suficiente para asignar una sesión a un determinado tema.

Para cada sesión se analizaron los conteos de coincidencias por tópico, considerando términos distintos del diccionario detectados al menos una vez, y se evaluó a partir de la magnitud de dichos conteos que reflejaban una dominancia temática clara y consistente con la evaluación manual.

Como resultado de este proceso iterativo de validación y ajuste, se establecieron los siguientes umbrales mínimos de activación:

- Economía ≥ 3 términos distintos
- Educación ≥ 3 términos distintos
- Salud ≥ 3 términos distintos
- Medioambiente ≥ 4 términos distintos
- Seguridad ≥ 3 términos distintos

Una sesión es considerada candidata de un tópico únicamente si el número de términos distintos detectados para dicho tópico supera su umbral correspondiente. En caso de que ningún tópico alcance su umbral, la sesión se clasifica como “Otros”, reflejando la ausencia de una predominancia temática clara.

El uso de términos distintos – en lugar de frecuencias brutas – responde a la necesidad de capturar diversidad léxica temática y evitar la sobre representación de tópicos debido a repeticiones retóricas de un mismo concepto dentro de sesiones extensas.

Cuando más de un tópico supera su umbral en una misma sesión, la asignación final se realiza en dos etapas. En primer lugar, se selecciona el tópico con mayor número de términos distintos detectados. En caso de empate, se aplica una jerarquía de desempate basada en la especificidad semántica del vocabulario de cada diccionario:

medioambiente > salud > seguridad > economía > educación

Esta jerarquía fue definida a partir de la revisión manual de la submuestra validada, considerando que ciertos tópicos presentan un vocabulario especializado cuya aparición suele indicar una focalización temática clara, mientras que otros contienen un mayor número de términos transversales que pueden aparecer en múltiples contextos legislativos sin constituir necesariamente el eje central del debate.

Con el diccionario maestro depurado y los criterios de decisión definidos a partir del proceso de validación descrito, el procedimiento fue aplicado al corpus completo de cada sesión del Senado. Cada sesión quedó asociada a un tópico dominante o a la categoría “Otros”, junto con los conteos de términos por tópico. Este enfoque constituye una aproximación híbrida que combina validación manual focalizada, reglas léxicas explícitas y expansión semántica automática, garantizando interpretabilidad y coherencia metodológica en la clasificación temática del discurso parlamentario.

4.2.6 Análisis de sentimiento basado en modelos Transformers

El análisis de sentimiento se incorporó como una segunda capa de interpretación del discurso parlamentario, orientada a estimar la orientación emocional de las intervenciones una vez definida su clasificación temática. Este análisis se aplicó exclusivamente a las sesiones que contaban con un tópico asignado, excluyendo aquellas clasificadas como “Otros”. Además, se utilizó el texto original de las sesiones parlamentarias, sin aplicar procedimientos adicionales de limpieza, eliminación de stopwords o lematización. Esta decisión responde a que los modelos de lenguaje basados en Transformers operan directamente sobre texto en lenguaje natural y han sido entrenados considerando la estructura sintáctica completa, así como el uso de conectores, marcadores discursivos y expresiones modales relevantes para la estimación de la carga emocional.

A diferencia del enfoque léxico utilizado en la clasificación temática, donde la normalización facilita la detección de patrones conceptuales, en el análisis de sentimiento la preservación del texto original permite mantener información pragmática y contextual - como negaciones, intensificadores y construcciones retóricas - que resulta clave para una interpretación adecuada del tono emocional del discurso parlamentario,

Dado que los modelos de lenguaje utilizados presentan una restricción en la longitud máxima de entrada de 512 tokens, cada sesión fue segmentada en bloques de 400 tokens, lo que permite asegurar compatibilidad técnica con los modelos Transformer, así como una aplicación homogénea del análisis de sentimiento a lo largo del corpus. El objetivo de esta división no es preservar unidades discursivas completas, si no obtener fragmentos comparables que permitan analizar la variación del tono emocional dentro de cada sesión.

Se evaluaron dos modelos de análisis de sentimientos preentrenados para español:

1. BETO fine-tuned para análisis de sentimiento, basado en la arquitectura BERT y ajustado específicamente para tareas de clasificación de polaridad en español. (Cañete et al., 2020; modelo finiteautomata/beto-sentiment-analysis)
2. RoBERTuito Sentiment, modelo de la librería pysentimiento entrenado principalmente sobre textos de carácter más informal y de uso general. (Pérez et al., 2021).

Antes de su aplicación al corpus completo, se realizó una validación metodológica con el objetivo de determinar cuál de los modelos capturaba de mejor manera la carga emocional del discurso legislativo. Para ello, se construyó una muestra de validación compuesta por 50 sesiones parlamentarias, correspondiente a 10 sesiones por cada uno de los cinco tópicos temáticos definidos, con el objetivo de asegurar diversidad temática en la evaluación.

Cada una de estas 50 sesiones fue segmentada en bloques de 400 tokens y clasificada con ambos modelos. A partir de este conjunto, se construyó una submuestra de validación compuesta por 60 bloques por modelo, distribuidos equitativamente en 20 bloques negativos, 20 neutrales y 20 positivos (120 bloques en total), según la clasificación automática de cada modelo. Esta submuestra fue sometida a una revisión manual, evaluando la correspondencia entre la predicción del modelo y el tono emocional observado en el texto.

Sobre esta base se calcularon métricas de desempeño, incluyendo accuracy y F-1 score, permitiendo comparar el comportamiento relativo de ambos modelos en el dominio legislativo. Los resultados mostraron que BETO fine-tuned presenta un desempeño sustancialmente superior, mientras que RoBERTuito exhibe confusiones sistemáticas, especialmente entre las categorías neutral y negativa. Este comportamiento se atribuye principalmente a las diferencias en los dominios lingüísticos de entrenamiento de cada modelo, siendo BETO más adecuado para textos formales, técnicos e institucionales como las sesiones del Senado.

En función de esta validación, BETO fine-tuned fue seleccionado como el modelo definitivo para el análisis de sentimiento del corpus. Una vez seleccionado el modelo, fue aplicado a la totalidad de las sesiones parlamentarias clasificadas temáticamente. Para cada sesión se obtuvieron predicciones de sentimiento a nivel de bloque, las cuales fueron posteriormente agregadas para construir indicadores a nivel de sesión.

En particular, se calcularon los siguientes indicadores:

- Moda y mediana del sentimiento, utilizada como clasificación principal de la sesión según la categoría más frecuente. Además, se calculó la media para evitar posibles desviaciones en sesiones más extensas.
- Promedio número del sentimiento, mediante la codificación de las categorías como -1 (negativo), 0 (neutral) y +1 (positivo), con el fin de capturar matices emocionales incluso cuando la moda resulta neutral.
- Distribución porcentual del sentimiento, expresada como la proporción de bloques negativos, neutrales y positivos dentro de cada sesión.
- Evolución temporal del sentimiento por deciles, que permite analizar cómo varía la carga emocional a lo largo del desarrollo interno de cada sesión, desde su inicio hasta su cierre.

Este conjunto de métricas permite caracterizar el tono emocional del discurso parlamentario no solo en términos de orientación dominante, si no también considerando su intensidad, composición y dinámica interna.

Con la clasificación temática del corpus y la definición del esquema de análisis de sentimiento descritos en este capítulo metodológico, la siguiente sección presenta los resultados obtenidos a partir de la aplicación de estos procedimientos al conjunto de sesiones del Senado. En particular, se analiza la distribución temática del discurso parlamentario, la orientación emocional asociadas a cada tópico y la evolución temporal del sentimiento, con el fin de contrastar las hipótesis planteadas y caracteriza patrones del discurso del Congreso Chileno.

5. Resultados

Con el objetivo de evaluar el desempeño inicial del enfoque de clasificación temática basado en diccionarios y evaluar el comportamiento preliminar del clasificador antes de su aplicación al corpus completo, se seleccionó una muestra de 350 sesiones del Senado. Sobre esta muestra se aplicó el procedimiento descrito en la sección 4.2, consistente en

identificar la presencia de términos pertenecientes al diccionario maestro previamente construido y asignar un tópico dominante a cada sesión en función de los criterios definidos.

5.1 Desempeño inicial del clasificador basado en diccionarios

La aplicación inicial del diccionario maestro sobre la muestra aleatoria de 350 sesiones parlamentarias mostró una distribución temática marcadamente desequilibrada entre los distintos tópicos considerados. La tabla N°2 presenta la distribución de sesiones clasificadas por tópico en esta etapa preliminar.

Tópico	Numero de sesiones clasificadas	Porcentaje del total
Economía	109	31,1%
Educación	20	5,7%
Medioambiente	38	10,9%
Salud	21	6,0%
Seguridad	99	28,3%
Otros	63	18,0%
Total	350	100%

Tabla N°2: Distribución de temáticas inicial

Los resultados evidencian una concentración desproporcionada de clasificaciones en los tópicos de economía y seguridad, los cuales en conjunto representan cerca del 60% de las sesiones clasificadas, este comportamiento no necesariamente refleja una predominancia real de estos temas en la agenda legislativa, si no que responde a características del vocabulario parlamentario y del funcionamiento inicial del clasificador léxico.

En particular, se identifican dos factores principales que explican esta distribución:

1. Alta transversalidad del vocabulario económico y de seguridad. Muchos términos asociados a estos tópicos (por ejemplo, recursos, presupuesto, sistema, control, estrategia) aparecen de manera recurrente en discusiones parlamentarias de naturaleza diversa, aun cuando no constituyen el eje central del debate. Esto

incrementa la probabilidad de activación de estos tópicos en sesiones cuyo foco principal pertenece a otras áreas.

2. Presencia de términos de bajo nivel de especificidad temática en los diccionarios iniciales. En esta etapa preliminar, algunos diccionarios contenían términos relativamente generales o contextualmente ambiguos, lo que favorecía su activación frente a diccionarios con vocabulario más especializado, como medioambiente o salud.

Este diagnóstico inicial puso de manifiesto la necesidad de refinar los diccionarios y los criterios de asignación temática, con el fin de mejorar la discriminación entre tópicos, reducir clasificaciones dominadas por vocabulario transversal y aumentar la coherencia temática de los resultados. Estos ajustes se abordarán en la siguiente sección.

5.2 Diagnóstico y refinamiento de errores

El análisis de la muestra inicial de 350 sesiones parlamentarias permitió identificar limitaciones relevantes en el comportamiento inicial del clasificador basado en diccionarios. En particular, la distribución temática presentada en la sección anterior evidenció una concentración excesiva de sesiones en los tópicos de economía y seguridad, lo que sugería una activación desproporcionada de estos diccionarios en contextos donde no necesariamente constituían el eje central del debate.

A partir de la revisión de los resultados de la clasificación inicial, se detectaron tres categorías principales de problemas:

- Términos transversales: palabras como “control”, “frontera”, “confianza” o “estrategia” aparecían en contextos legislativos diversos, activando tópicos específicos sin representar necesariamente el eje central del debate.
- Expresiones excesivamente generales: términos como “recurso”, “riesgo”, “inversión” o “administrativo” se encontraban distribuidos en múltiples áreas de discusión parlamentaria, lo que favorecía asignaciones temáticas poco precisas.
- Ambigüedades derivadas de la lematización, en algunos casos, los procesos de normalización y lematización redujeron expresiones específicas a formas más generales que coincidían con vocabulario frecuente en ámbitos administrativos. Por ejemplo, expresiones compuestas como “estructuras críticas” fueron lematizadas a formas como “estructura crítico”, lo que favorecía su activación en contextos no vinculados a seguridad. Este tipo de reducción eliminó información contextual relevante, generando activaciones no representativas y ruido semántico en la clasificación temática.

Estos fenómenos se observaron con mayor intensidad en los diccionarios de economía y seguridad durante esta primera etapa, cuyos vocabularios iniciales incluían una mayor proporción de términos de uso transversal en el discurso parlamentario. Esto explica la sobrerrepresentación de estos tópicos en la clasificación inicial.

A partir de este diagnóstico, se implementó un proceso iterativo de depuración del diccionario maestro. Este proceso incluyó la eliminación manual de términos genéricos, ambiguos o poco específicos, priorizando expresiones con mayor anclaje conceptual en cada área temática.

Una vez aplicado este refinamiento inicial, el diccionario ajustado fue utilizado para clasificar nuevamente la misma muestra de 350 sesiones parlamentarias. El resultado de esta segunda clasificación mostró una redistribución temática más equilibrada entre los distintos tópicos, reduciendo la concentración observada previamente en economía y seguridad.

Tópico	Numero de sesiones clasificadas	Porcentaje del total
Economía	92	26,3%
Educación	77	22,0%
Medioambiente	86	24,6%
Salud	45	12,8%
Seguridad	48	13,7%
Otros	2	0,6%
Total	350	100%

Tabla N°3 Resultado de clasificación con diccionario refinado

Con el fin de evaluar más detalladamente la calidad de esta clasificación evaluar más detalladamente la calidad de esta clasificación refinada, se seleccionó posteriormente una submuestra de 52 sesiones para revisión manual, compuesta por 10 sesiones por cada uno de los cinco tópicos temáticas y 2 sesiones clasificadas como “Otros”. Esta submuestra permitió contrastar la asignación automática con la lectura directa del contenido de las sesiones, identificando errores residuales y patrones de confusión específicos en cada diccionario.

La revisión manual de esta submuestra reveló que, si bien la clasificación había mejorado, persistían activaciones no representativas asociadas principalmente a ciertos términos conflictivos en los diccionarios de educación y medioambiente. En estos casos, vocablos relacionados con ámbitos profesiones, administrativos o productivos aparecían en contextos no educativos o no ambientales generando asignaciones temáticas incorrectas.

A partir de esta revisión focalizada, se realizó una última depuración del diccionario maestro, eliminando conjuntos específicos de términos problemáticos y ajustando la composición léxica de los diccionarios afectados. Este proceso no tuvo como objetivos maximizar el tamaño de los diccionarios, sino armonizar su contenido y asegurar que las expresiones retenidas presentaran una alta especificidad temática dentro del discurso parlamentario.

El efecto acumulado de este proceso iterativo de expansión semántica, validación empírica y depuración manual se resume en la siguiente comparación de los tamaños de los diccionarios en sus distintas etapas.

Tópico	Diccionario inicial	Diccionario expandido	Diccionario refinado
Economía	55	69	74
Educación	69	107	86
Medioambiente	77	110	73
Salud	94	112	103
Seguridad	91	129	71
Total	386	527	407

Tabla N°4 comparación tamaño original, expandido y refinado de diccionarios

La revisión manual de la submuestra de 52 sesiones permitió identificar patrones que fundamentan las decisiones metodológicas adoptadas en la definición final del esquema de clasificación temática. En particular, se observó que las asignaciones incorrectas se concentraban mayoritariamente en sesiones donde el tópico dominante se activaba a partir de uno o dos términos aislados, generalmente de carácter transversal o no representativo, mientras que las clasificaciones correctas tendían a presentar una mayor diversidad léxica temática. Este comportamiento fue especialmente evidente en los tópicos de economía y seguridad, cuyos diccionarios iniciales incluían vocabulario de uso frecuente en múltiples contextos legislativos.

A partir de esta evidencia, se establecieron umbrales mínimos diferenciados por tópicos antes de asignar una sesión a un eje temático. Asimismo, el análisis de casos con múltiples tópicos activados mostró que los empates eran relativamente poco frecuentes en la submuestra analizada y que podían resolverse de manera consistente mediante una jerarquía basa en la especificidad semántica de los diccionarios, privilegiando aquellos cuya aparición indica una focalización temática más clara. En conjunto, estos resultados respaldan el carácter iterativo del proceso de refinamiento y justifican que los criterios finales de clasificación no responden a decisiones arbitrarias, si no a patrones observados sistemáticamente durante la validación manual.

Tópico	Clasificaciones Correctas	Promedio de términos	Mínimo
Economía	9	4,8	3
Educación	7	4,1	3
Medioambiente	7	5,6	4
Salud	9	5,0	3
Seguridad	10	5,8	3

Tabla N°5 Diversidad léxica en sesiones correctamente clasificadas

La Tabla N°5 resume la diversidad léxica observada en las sesiones correctamente clasificadas durante la revisión manual, mostrando que estas presentan sistemáticamente un número de términos igual o superior a los umbrales finalmente adoptados.

En conjunto, estos resultados confirman que el refinamiento progresivo del diccionario maestro permitió mejorar sustancialmente la coherencia temática de la clasificación, reduciendo activaciones no representativas y fortaleciendo la correspondencia entre los tópicos asignados y el contenido efectivo de las sesiones. Este proceso iterativo constituye una etapa clave para garantizar la validez conceptual de la clasificación temática antes de su aplicación definitiva al corpus completo del Senado.

5.3 Distribución temática del discurso parlamentario

Una vez definido el esquema de clasificación temática y validados los criterios de decisión descritos en las secciones 4.2.5 y 5.2, se procedió a aplicar el clasificador temático al corpus completo de sesiones del Senado, compuesto por 1.674 sesiones correspondientes al periodo 2010-2025. Este análisis permite caracterizar la agenda temática del debate parlamentario chileno desde una perspectiva agregada y temporal, identificando los ejes de la política pública que concentran mayor atención legislativa.

La tabla N°6 presenta la distribución global de sesiones clasificadas por tópico, considerando el diccionario refinado, los umbrales mínimos de activación por tópico y la jerarquía de desempate definidos en la metodología. La categoría “Otros” agrupa aquellas sesiones en las que no se detectó una predominancia temática clara según los criterios establecidos.

Tópico	Clasificados	Porcentaje
Economía	377	22,52%
Educación	292	17,44%
Medioambiente	151	9,02%
Salud	321	19,18%
Seguridad	357	21,33%
Otros	176	10,51%
Total	1.674	100%

Tabla N°6 distribución temática global del corpus

Los resultados muestran que el discurso parlamentario se concentra principalmente en los tópicos de economía (22,5%), seguridad (21,3%) y salud (19,2%), los cuales en conjunto representan más del 60% de total de sesiones analizadas. En un segundo nivel se ubican educación (17,4%) y medioambiente (9,0%), mientras que aproximadamente un 10,5% de las sesiones no presenta un eje temático dominando bajo el esquema de clasificación aplicado.

Esta distribución resulta coherente con la naturaleza del debate legislativo, donde materias económicas, de seguridad y sanitarias tienden a ocupar un lugar central en la agenda parlamentaria, especialmente en contextos de crisis, reformas estructurales o discusión presupuestaria. Asimismo, la presencia de una proporción acotada de sesiones clasificadas como “Otros” refleja la existencia de debates procedimentales, administrativos o de carácter transversal que no se alinean claramente con los cinco ejes temáticos definidos. Con el fin de analizar como varía la agenda temática del Senado a lo largo del tiempo, se examinó la distribución porcentual de sesiones por tópico para cada año del periodo de estudio. Para este análisis se excluyó la categoría “Otros”, concentrándose únicamente en las sesiones con clasificación temática definida, lo que permite comprar de manera más clara la dinámica relativa entre los distintos ejes de política pública.

La figura N°1 muestra la evolución temporal de la cantidad de sesiones asociadas a cada tópico entre 2010 y 2025.

El gráfico evidencia que la importancia relativa de los distintos tópicos no es estática, más bien que presenta variaciones relevantes a lo largo del tiempo. En los primeros años del periodo analizado, el debate parlamentario aparece fuertemente dominado por materias económicas, mientras que en otros momentos se observa un aumento significativo de tópicos como salud y seguridad, particularmente en años asociados a crisis sanitarias, conflictos de orden público o estados de excepción.

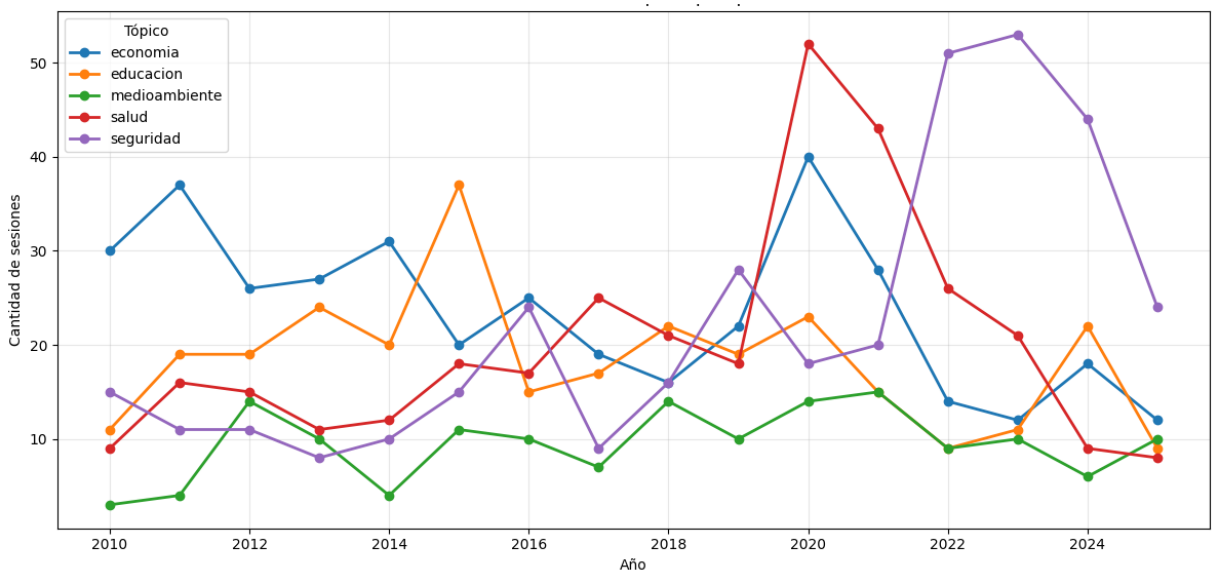


Figura N°1 evolución temporal de sesiones por tópico y por año

Asimismo, el tópico medioambiente, si bien mantiene una participación relativa menor en términos globales, muestra incrementos puntuales en determinados años, lo que sugiere la emergencia de agendas ambientales específicas vinculadas a eventos críticos, reformas regulatorias o compromisos internacionales. Por su parte educación presenta una participación relativamente estable a lo largo del período, aunque con incrementos pronunciados en años específicos, destacando particularmente 2015, donde alcanza su mayor proporción relativa de sesiones. Este aumento coincide con un ciclo de intensa discusión legislativa en torno a reformas estructurales del sistema educativo y debates presupuestarios asociados, lo que sugiere una activación temática coyuntural más que una tendencia sostenida en el tiempo. Durante este periodo, los picos de presencia temática observados por tópico se explican principalmente por coyunturas político-institucionales específicas que cambian la agenda legislativa. En economía, se observan incrementos relevantes en los años 2011 y 2020. El incremento en 2011 se asocia a un ciclo intenso de discusión legislativa en torno a política tributaria, recaudación fiscal y financiamiento del gasto público, estrechamente vinculado a las demandas sociales emergentes durante el movimiento estudiantil de este año. En 2020, en cambio, el aumento se vincula directamente a la crisis económica derivada de la pandemia COVID-19, que concentro la discusión en medidas de emergencia como apoyo a empleo y empresas, endeudamiento público, entre otros.

En educación, el pico de 2015 se vincula con la reforma educacional del segundo gobierno de Michelle Bachelet y hitos legislativos como la ley de inclusión escolar (Ley N°20.845, 2015), además del debate sobre gratuidad y la regulación de la educación superior.

En salud, el máximo de 2020 responde directamente a la pandemia, con discusión concentrada en medidas sanitarias, gestión hospitalaria, licencias y protección del personal.

En medioambiente, los picos de 2012, 2018, 2020 y 2021 reflejan, una activación temática vinculada principalmente a crisis ambientales concretas. El incremento en 2012 coincide con un periodo de alta conflictividad socioambiental asociada a proyectos energéticos y de infraestructura, como HidroAysén. Los aumentos observados en 2018, 2020 y especialmente en 2021 se explican por la creciente centralidad de la crisis hídrica en el debate parlamentario, con discusiones recurrentes sobre escasez de agua, abastecimiento para consumo humano, decretos de escasez hídrica, uso de camiones aljibe y gestión de los derechos de aprovechamiento de aguas, lo que refuerza el carácter reactivo y aplicado del discurso ambiental.

Finalmente, en seguridad, los incrementos de 2019, 2022 y 2023 se asocian, primero, al estallido social iniciado en octubre de 2019, con debates sobre orden público, violencia, derechos humanos y estados de excepción, posteriormente, a una agenda legislativa centrada en el combate al crimen organizado y la delincuencia, junto con discusiones sobre

fortalecimiento policial, reformas al sistema penal y la aplicación reiterada de estados de excepción en la macrozona sur.

La caracterización de la distribución temática global y su evolución temporal permite establecer un marco descriptivo sólido sobre qué se discute en el Senado chileno y cómo estas prioridades varían a lo largo del tiempo. Estos resultados constituyen la base necesaria para el análisis presentado en la siguiente sección, donde se examina la orientación emocional del discurso parlamentario y su relación con los distintos tópicos temáticos identificados.

5.4 Validación y selección del modelo de análisis de sentimiento

Con el fin de abordar la estimación de la orientación emocional del discurso parlamentario mediante modelos de lenguaje y seleccionar aquel con mejor desempeño para el dominio legislativo, se realizó una validación comparativa entre dos modelos Transformers preentrenados para español: BETO fine-tuned (Cañete et al., 2020; modelo finiteautomata/beto-sentiment-analysis) y RoBERTuito Sentiment (Pérez et al., 2021). La validación se realizó sobre una muestra de 50 sesiones parlamentarias, seleccionadas con criterio de diversidad temática (10 sesiones por cada uno de los cinco tópicos definidos). Considerando la restricción técnica de la longitud máxima de entrada (512 tokens), cada sesión fue segmentada en bloques de aproximadamente 400 tokens, garantizando una aplicación homogénea del análisis entre sesiones y entre modelos.

A partir de las predicciones obtenidas por ambos modelos sobre los bloques de la muestra, se construyó una submuestra balanceada de validación manual compuesta por 60 bloques por modelo (20 NEG, 20 NEU y 20 POS según la clasificación), totalizando 120 bloques revisados manualmente. Esta estrategia permitió evaluar el rendimiento con equilibrio entre clases y medir directamente la capacidad de los modelos para capturar carga emocional en un registro institucional, formal y técnico como el legislativo.

Con el fin de caracterizar la estructura de la muestra utilizada para la validación comparativa de los modelos de análisis de sentimiento, se analizó el número de bloques generados por sesión tras la segmentación del texto en fragmentos de aproximadamente 400 tokens.

La Tabla N°7 presenta estadísticas descriptivas del volumen de bloques por sesión para cada modelo, calculadas sobre la muestra de 50 sesiones parlamentarias. En promedio, las

sesiones se dividieron en 94 bloques al utilizar BETO fine-tuned y en 87 bloques al utilizar RoBERTuito, con una variabilidad considerable entre sesiones, reflejando la heterogeneidad en la extensión de los debates parlamentarios.

Estadística Descriptiva	Bloques por sesión Beto fine-tuned	Bloques por sesión RoBERTuito
Frecuencia	50	50
Promedio	94,38	87,28
Desviación estándar	54,21	49,85
Mínimo	23	22
25%	64,5	59,5
50%	82	77,5
75%	107,75	98
Máximo	349	321

Tabla N°7 Estadísticos descriptivos del número de bloques por sesión (muestra de 50 sesiones; BETO vs RoBERTuito)

Una vez caracterizada la estructura de la muestra y el volumen de fragmentos analizados por sesión, el siguiente paso consiste en examinar el comportamiento de los modelos de análisis de sentimiento seleccionados, considerando su adecuación al dominio lingüístico del discurso parlamentario.

Los modelos evaluados fueron:

- BETO fine-tuned para análisis de sentimiento, basado en la arquitectura BERT para español y ajustado específicamente para tareas de clasificación de polaridad en registros formales.
- RoBERTuito Sentiment, disponible en la librería pysentimiento, entrenado mayoritariamente sobre textos de uso general y registros menos formales.

Dado el estilo del discurso parlamentario - frecuentemente normativo, argumentativo y con alta densidad técnica - resulta esperable anticipar que un modelo entrenado sobre español formal presentara un mejor desempeño relativo. Esta expectativa sirve como referencia para interpretar los resultados de la revisión manual que se presentan a continuación.

La evaluación de los modelos de análisis de sentimiento se realizó sobre una submuestra de validación de 60 bloques por modelo (20 bloques negativos, 20 bloques positivos y 20 bloques neutrales), contrastando la clasificación automática del modelo con una clasificación manual. Los resultados se presentan a continuación:

La Tabla N°8 presenta la matriz de confusión correspondiente al modelo BETO fine-tuned. Los resultados muestran una clasificación altamente consistente en las tres categorías. En particular, se observa una separación casi perfecta entre las clases negativas, neutral y positiva, con un único error de clasificación en la categoría positivo.

	Predicho Negativo	Predicho Neutral	Predicho Positivo
Real Negativo	20	0	1
Real Neutral	0	20	0
Real Positivo	0	0	19

Tabla N°8 Matriz de confusión Modelo BETO fine-tuned

Este patrón se refleja en métricas de desempeño elevadas, con un nivel de exactitud cercano al máximo teórico, lo que indica que el modelo captura de manera adecuada tanto valoraciones explícitas como críticas expresadas mediante un registro institucional.

La Tabla N°9 muestra la matriz de confusión correspondiente al modelo RoBERTuito Sentiment. A diferencia del modelo BETO, este modelo presenta una mayor dispersión de errores, particularmente en la distinción entre las categorías Negativo y Neutral. Se observa que una proporción relevante de fragmentos evaluados manualmente como negativos es clasificada como neutral, así como confusiones cruzadas en menor medida entre las demás clases.

	Predicho Negativo	Predicho Neutral	Predicho Positivo
Real Negativo	15	7	0
Real Neutral	3	11	1
Real Positivo	3	2	19

Tabla N°9 Matriz de confusión Modelo RoBERTuito Sentiment

Este comportamiento sugiere una menor estabilidad clasificatoria en contextos donde la evaluación negativa se expresa de manera formal o técnica, sin marcadores emocionales explícitos.

La comparación directa del desempeño de ambos modelos se resume en la tabla N°10, que presenta los valores de accuracy, precisión, recall y F-1 score. En esta tabla se aprecia una diferencia a favor de BETO fine-tuned, que alcanzan valores cercanos a la clasificación perfecta en todas las métricas, mientras que RoBERTuito muestra una reducción consistente en su desempeño.

Métrica	Modelo BETO	Modelo RoBERTuito
Accuracy	0,983	0,750
Precision	0,983	0,750
Recall	0,984	0,753
F-1 Score	0,983	0,741

Tabla N°10 Métricas comparativas de desempeño – BETO vs RoBERTuito

Cabe señalar que los valores de las métricas resultan numéricamente similares, lo cual se explica por el diseño balanceado de la submuestra de validación (20 bloques por clase) y por la distribución relativamente homogénea de los errores entre categorías.

En conjunto, estos resultados cuantitativos confirman que BETO fine-tuned ofrece una mayor precisión y estabilidad en la clasificación del sentimiento en el discurso parlamentario, constituyendo una base empírica sólida para su selección como modelo definitivo en el análisis del corpus completo.

5.5 Análisis del sentimiento del discurso parlamentario

Una vez validado y seleccionado el modelo BETO fine-tuned como herramienta definitiva para el análisis de sentimiento, esta sección presenta los resultados obtenidos tras su aplicación al corpus completo de sesiones parlamentarias clasificadas temáticamente. Siguiendo lo definido en la metodología, se excluyeron del análisis aquellas sesiones sin tópico asignado “*Otros*”, resultando en un total de 1.498 sesiones con tópico asignado, segmentadas en 140.977 bloques de 400 tokens.

El análisis que se presenta a continuación examina la distribución global del sentimiento, su variación por tópico, así como su dinámica temporal, tanto a lo largo del tiempo legislativo como dentro de la estructura interna de las sesiones.

El primer nivel de análisis considera la distribución agregada del sentimiento a nivel de bloques, lo que permite caracterizar el tono general del discurso parlamentario en el periodo estudiado.

Sentimiento	Cantidad de bloques	Porcentaje (%)
Negativo	44.822	31,79%
Neutral	85.932	60,95%
Positivo	10.223	7,26%
Total	140.977	100%

Tabla N°11 Distribución global del sentimiento a nivel de bloques

La tabla N°11 muestra un claro predominio neutral, que concentra aproximadamente el 61% de los bloques, seguido por una proporción relevante de bloques negativos 31,79%, mientras que el sentimiento positivo aparece de manera mucho más acotada 7,26%. Este patrón es consistente con la naturaleza procedimental, normativa y técnica del debate legislativo, donde gran parte de las intervenciones cumplen funciones informativas o argumentativas sin una carga emocional explícita. El patrón negativo se asocia principalmente a formas de negatividad evaluativa e institucional, más que a expresiones insultantes. En particular, predominan:

- Criticas a políticas públicas y diagnósticos de insuficiencia o mal funcionamiento
- Marcos de preocupación o urgencia vinculados a crisis sociales, sanitarias o de seguridad

- Confrontación política mediante impugnación del adversario o denuncia de inconsistencias.

Además, se observan bloques de intensidad retórica (condena moral explícita) y bloques negativos vinculados a memoria y derechos humanos, donde la carga emocional surge del contenido de reparación y justicia más que del conflicto partidario.

Con el fin de evitar posibles desviaciones que podrían introducir las sesiones extensas (con mayor número de bloques), el análisis se trasladó al nivel de sesión, calculado el sentimiento promedio y mediana por sesión por tópico.

Tópico	Media del sentimiento	Mediana del sentimiento	Cantidad de sesiones
Economía	-0,254	-0,252	377
Educación	-0,203	-0,222	292
Medioambiente	-0,210	-0,208	151
Salud	-0,227	-0,230	321
Seguridad	-0,224	-0,226	357

Tabla N°12 Sentimiento promedio y mediana por sesión por tópico

Los valores presentados en la Tabla N°12 indican que todos los tópicos presentan, en promedio, un sentimiento negativo leve, aunque con diferencias relevantes. El discurso en economía exhibe el promedio más negativo, mientras que educación y medioambiente muestran tonos comparativamente menos negativos.

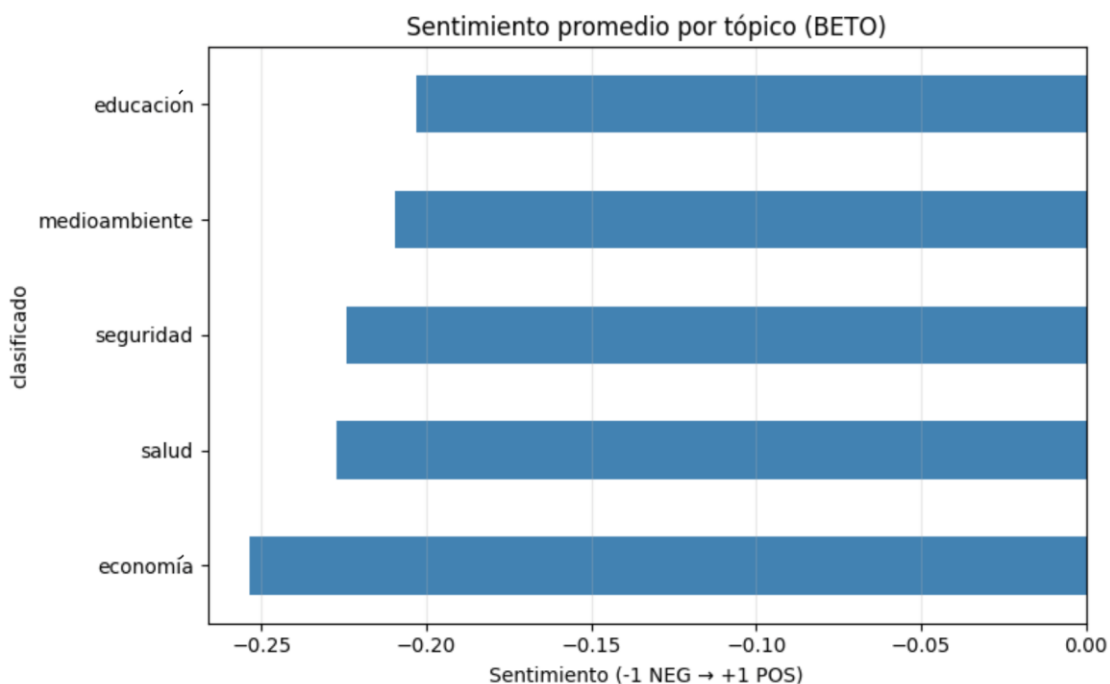


Gráfico N°2 Sentimiento promedio por tópico

El gráfico N°2 visualiza estas diferencias, mostrando diferencias en la magnitud del valor promedio entre los distintos ejes temáticos. En términos generales, ningún tópico alcanza valores positivos promedio, y todos presentan un tono negativo, aunque con leves variaciones en su intensidad. Estas diferencias deben interpretarse como una caracterización descriptiva del tono emocional asociado a cada ámbito del debate parlamentario, lo que sugiere la existencia de estilos discursivos diferenciados según el tópico tratado.

Para complementar el análisis a nivel de sesión, se examinó la composición porcentual del sentimiento de cada tópico.

Tópico	Negativo	Neutral	Positivo
Economía	34,06%	58,95%	6,99%
Educación	31,82%	60,62%	7,56%
Medioambiente	29,71%	62,26%	8,03%
Salud	31,69%	61,33%	6,98%

Seguridad	30,41%	62,40%	7,19%
-----------	--------	--------	-------

Tabla N°13 Distribución porcentual del sentimiento por tópico.

La Tabla N°13 muestra que, si bien el sentimiento neutral domina en todos los tópicos, el sentimiento negativo representa de manera consistente alrededor de un 30% de los bloques analizados en cada eje temático. El tópico de medioambiente concentra el mayor porcentaje de bloques positivos (8,03%); sin embargo, sigue siendo claramente minoritaria en comparación con las categorías neutral y negativa.

El Gráfico N°4 muestra la evolución temporal del sentimiento promedio desagregado por tópico entre 2010 y 2025.

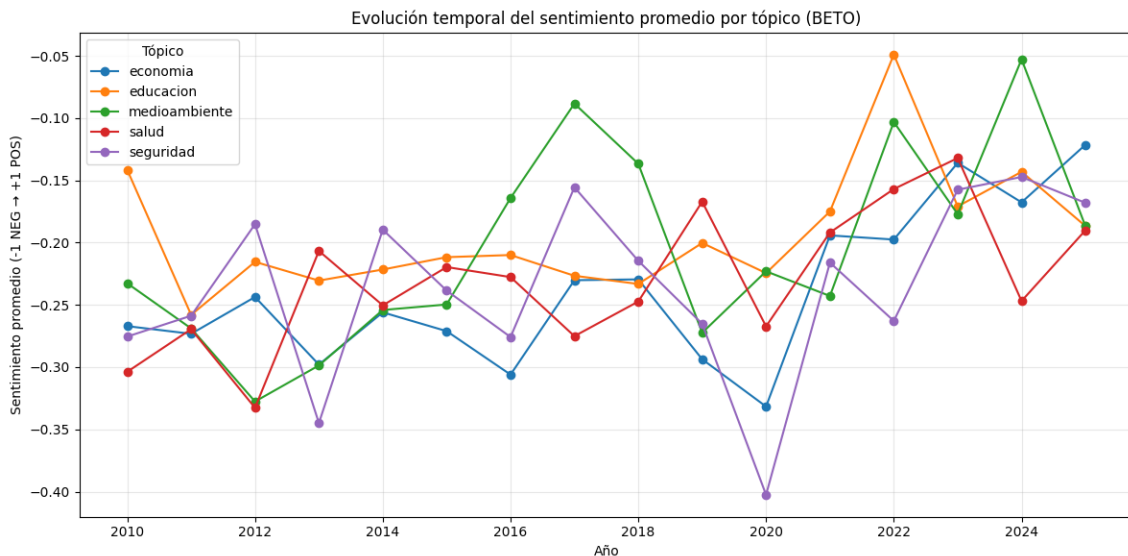


Gráfico N°4 Evolución temporal del sentimiento promedio por tópico (2010-2025)

Si bien todos los ejes temáticos presentan fluctuaciones a lo largo del período, destaca particularmente el comportamiento del tópico de seguridad, que alcanza su valor más negativo en el año 2020, situándose por debajo del resto de los tópicos analizados. Este deterioro coincide con un período marcado por una alta conflictividad social y una fuerte presión institucional, lo que se refleja en un discurso parlamentario con mayor carga crítica y negativa. A partir de 2021, el sentimiento asociado a seguridad muestra una recuperación progresiva, alineándose nuevamente con los niveles observados en otros tópicos. En contraste, educación y medioambiente exhiben variaciones más moderadas, mientras que economía y salud presentan trayectorias relativamente más estables en el largo plazo.

Además del análisis agregado, se examinó como varía el sentimiento a lo largo de desarrollo interno de las sesiones, segmentándolos en deciles (inicio al cierre)

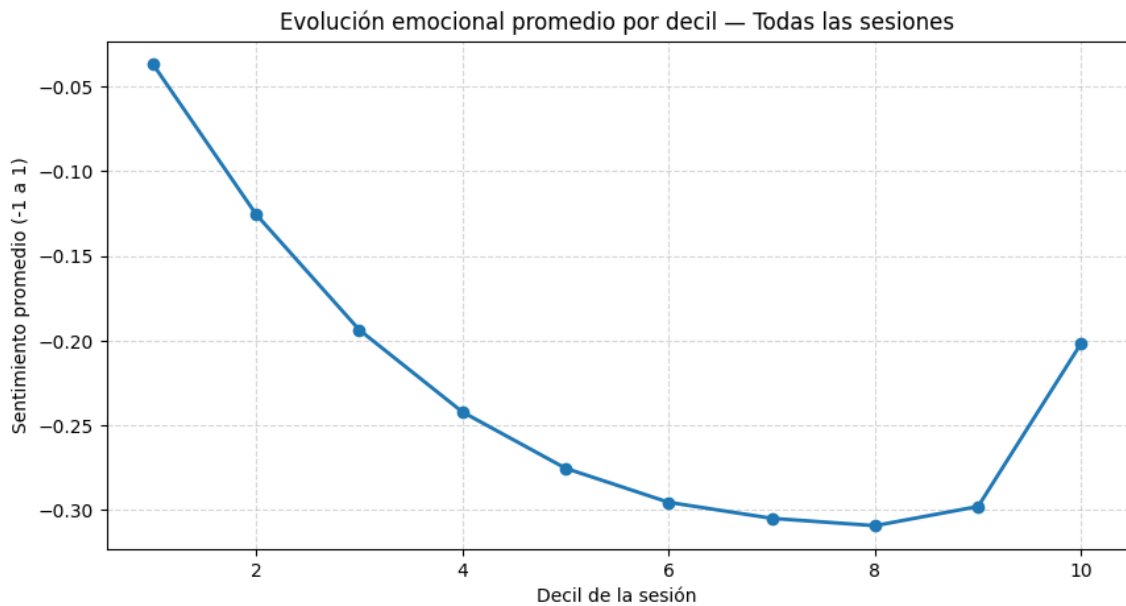


Gráfico N°5 Evolución del sentimiento promedio por decil de todas las sesiones

El Gráfico N°5 muestra un patrón sistémico, el sentimiento es menos negativo al inicio de las sesiones, se intensifica en los deciles intermedios, y tiende a moderarse hacia el cierre. Se podría pensar que este comportamiento es consistente con la estructura típica del debate parlamentario, donde los primeros segmentos suelen concentrar formalidades, cuenta y orden del día, mientras que el núcleo deliberativo se ubica en la parte central.

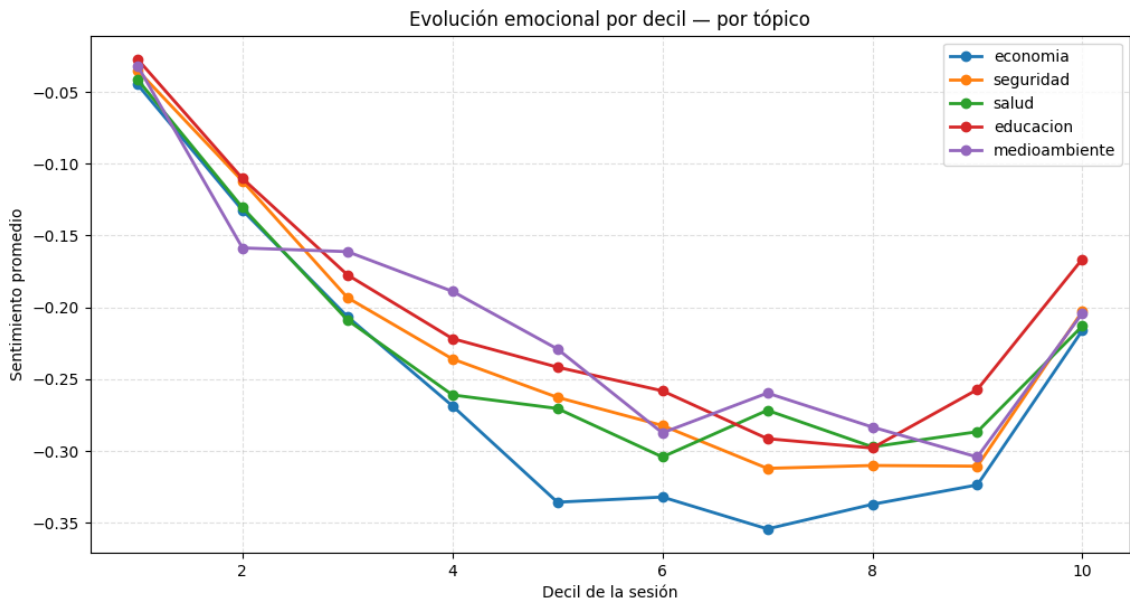


Gráfico N°6 Evolución del sentimiento por decil por tópico

Al separar la evolución emocional por tópico (Gráfico N°6), se observa que el patrón general de descenso del sentimiento se mantiene hacia los deciles centrales, aunque con diferencias de intensidad. En particular, economía presenta una caída más pronunciada hacia valores negativos en los deciles centrales, alcanzando su punto mínimo alrededor del séptimo decil. Seguridad y salud exhiben trayectorias similares, con descensos marcados y una recuperación gradual hacia el cierre, mientras que educación y medioambiente muestran trayectorias relativamente más estables. En todos los tópicos se observa un repunte del sentimiento en el último decil, lo que indica que las fases finales de las sesiones tienden a incorporar un tono menos negativo, independiente del contenido temático tratado.

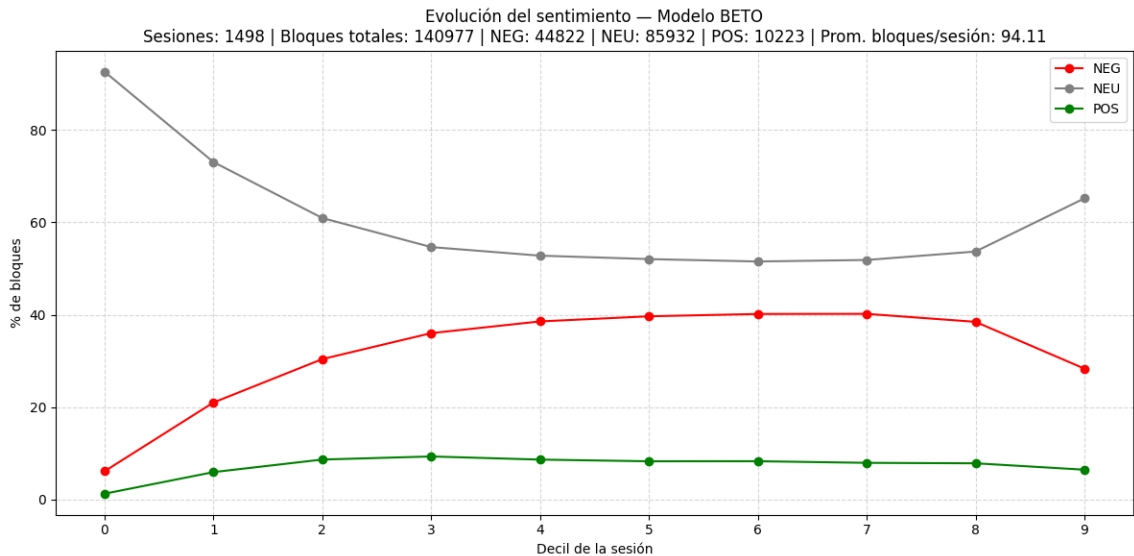


Gráfico N°7 Evolución del sentimiento por decil (NEG / NEU / POS)

El Gráfico N°7, resulta especialmente informativo para comprender la dinámica interna del discurso parlamentario a lo largo de las sesiones. En los primeros deciles se observa una clara dominancia de bloques clasificados como neutrales, lo que sugiere que las fases iniciales de las sesiones están mayoritariamente ocupadas por procedimientos formales, como la asistencia, el orden del día, la lectura de antecedentes y otros actos administrativos con baja carga emocional. A medida que avanza la sesión, los deciles intermedios muestran un aumento sostenido en la proporción de bloques negativos, alcanzando su punto máximo aproximadamente entre los deciles quinto y séptimo, coincidiendo con el momento en que se desarrollan los debates y la confrontación argumentativa entre posiciones políticas. Hacia los deciles finales, las proporciones de bloques negativos comienza a descender, mientras que los bloques neutrales vuelven a incrementarse, lo que sugiere una transición hacia etapas de cierre, votaciones o formalización de acuerdos. Este patrón refuerza la idea de que la carga emocional del discurso parlamentario no se distribuye de manera homogénea dentro de las sesiones, sino que responde a una estructura institucional recurrente.

Finalmente, se analizó la evolución temporal del sentimiento promedio del discurso parlamentario a lo largo del periodo 2010-2025.

El Gráfico N°8 muestra una variación temporal clara en la carga emocional del discurso parlamentario. Durante todo el periodo analizado, el sentimiento promedio se mantiene en valores negativos relativamente estables, reflejando un tono crítico persistente en el debate legislativo. Esta tendencia alcanza su punto más bajo en el año 2020, donde se observa el valor más negativo de toda la serie, lo que resulta coherente con un contexto

de alta tensión política e institucional, marcado simultáneamente por las consecuencias del estallido social iniciado a fines de 2019 y por la irrupción de la pandemia COVID-19.

A partir de 2021, se aprecia un cambio progresivo hacia valores menos negativos, con una mejora sostenida del sentimiento promedio hasta 2024, lo que sugiere una moderación relativa del tono emocional del discurso en los años posteriores al periodo de mayor conflictividad. Aunque el sentimiento promedio permanece en terreno negativo, la magnitud de esta negatividad disminuye de forma consistente, indicando una reconfiguración del clima discursivo más que un cambio a un tono predominantemente positivo.

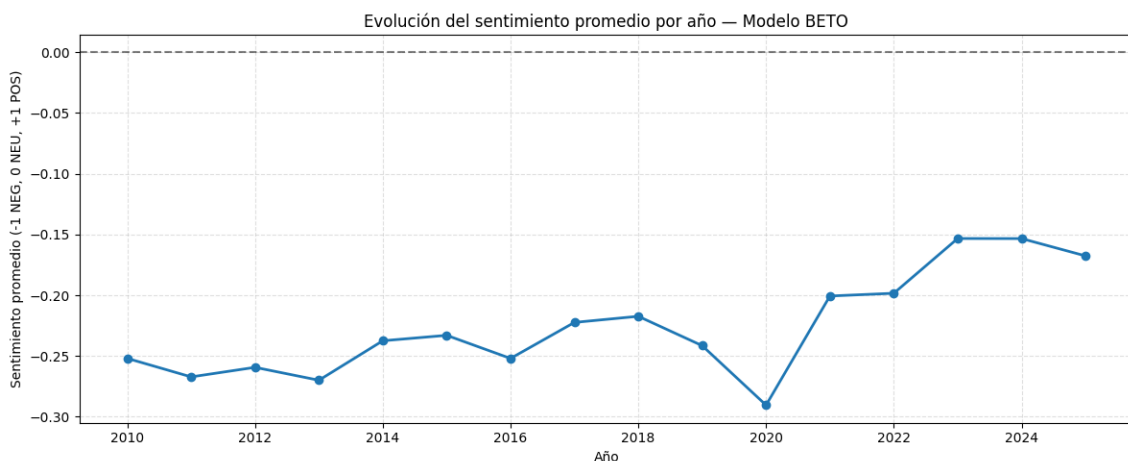


Gráfico N°8 Evolución temporal del sentimiento por año

En conjunto, este análisis temporal refuerza la hipótesis que la carga emocional del discurso parlamentario no es estática, si no que responde de manera sensible a contextos políticos, sociales e institucionales específicos, y que estos cambios pueden ser capturados mediante modelos de análisis de sentimiento aplicados de forma sistemática al corpus legislativo.

Los análisis presentados permiten confirmar que el discurso parlamentario chileno exhibe patrones emocionales sistemáticos, diferenciados tanto por tópico, como por estructura interna de las sesiones y por contexto temporal. La combinación del enfoque temático con el análisis de sentimiento revela que, junto a una agenda discursiva relativamente estable, la carga emocional varía – aunque predominantemente neutral - en función del contenido tratado, del momento del debate y de coyunturas políticas específicas. Estos hallazgos respaldan las hipótesis planteadas, mostrando que el tono emocional constituye una dimensión relevante y observable del discurso legislativo, y que su estudio aporta

información complementaria para comprender estilos comunicativos, prioridades políticas y dinámicas institucionales del Congreso Chileno.

6. Conclusiones

El presente estudio tuvo como objetivo analizar el discurso parlamentario chileno desde una perspectiva temática y emocional, utilizando técnicas de procesamiento de lenguaje natural aplicadas a un corpus de gran escala compuesto por sesiones del Senado entre los años 2010 y 2025. Los resultados obtenidos permiten confirmar la hipótesis central del proyecto: el discurso parlamentario no solo se estructura en torno a ejes temáticos diferenciados, sino que también presenta variaciones sistemáticas en su carga emocional, las cuales reflejan dinámicas políticas, estilos comunicativos e influencias contextuales de carácter institucional y social.

En primer lugar, se demostró que la estructuración temática del discurso parlamentario puede ser identificada con un nivel razonable de precisión mediante un enfoque basado en diccionarios especializados. La construcción progresiva del diccionario temático – a partir de semillas conceptuales, expansión semántica mediante modelos de embeddings y sucesivas instancias de revisión manual – permitió capturar patrones temáticos coherentes y estables en el corpus. El proceso de refinamiento, validado a través de una submuestra revisada manualmente, redujo activaciones falsas y fortaleció la correspondencia entre los tópicos asignados y el contenido sustantivo de las sesiones.

En segundo lugar, la clasificación temática del corpus completo evidenció una agenda discursiva claramente diferenciada entre los 5 ejes temáticos definidos, tanto en términos de volumen como de evolución temporal. Estos resultados confirman que los tópicos seleccionados constituyen dimensiones centrales del debate parlamentario y que su presencia varía de manera consistente a lo largo del periodo analizado. Este proceso fue acompañado por instancias de validación manual que permitieron evaluar la coherencia y precisión de la asignación temática, otorgando respaldo empírico al enfoque adoptado.

En tercer lugar, la incorporación del análisis de sentimiento permitió caracterizar la dimensión emocional del discurso parlamentario de forma sistemática. Los resultados muestran un claro predominio del sentimiento neutral, coherente con el carácter procedimental, normativo y argumentativo del debate legislativo. No obstante, se identificó también una presencia persistente de sentimiento negativo en todos los tópicos analizados, lo que sugiere que la crítica, la evaluación problematizadora y la confrontación argumentativa constituyen componentes estructurales del discurso parlamentario chileno. El sentimiento positivo, en contraste, aparece de manera marginal, reforzando la idea de un estilo discursivo predominantemente crítico, más que celebratorio.

Adicionalmente, el análisis temporal del sentimiento evidenció que la carga emocional del discurso no es estática. Se observaron variaciones significativas tanto a lo largo del tiempo legislativo analizado, como al interior de las sesiones. En particular, el año 2020 destaca como un punto de inflexión, con niveles de negatividad notablemente superiores al resto del periodo, lo que resulta coherente con un contexto marcado por alta conflictividad social e institucional. Asimismo, el análisis por deciles mostró patrones recurrentes en la estructura interna de las sesiones, con fases iniciales dominadas por neutralidad y un aumento de la negatividad en los segmentos centrales del debate.

En conjunto, estos hallazgos permiten afirmar que la combinación de análisis temático y emocional constituye una herramienta eficaz para capturar dimensiones del discurso parlamentario chileno, aportando evidencia que respalda las hipótesis planteadas y contribuye a una caracterización más profunda del estilo discursivo del Congreso chileno.

Este trabajo aporta una estrategia metodológica integrada para el análisis del discurso parlamentario a gran escala, combinando clasificación temática basada en diccionarios con modelos de análisis de sentimiento basados en Transformers. En particular, se propone un enfoque iterativo de construcción y refinamiento de diccionarios temáticos, que articula expansión automática mediante embeddings y validación manual, permitiendo adaptar el vocabulario al registro específico del debate legislativo.

Asimismo, el estudio demuestra la importancia de validar empíricamente los modelos de análisis de sentimiento en función de su adecuación al dominio lingüístico, evidenciando que modelos entrenados sobre español formal presentan un desempeño superior en contextos institucionales. Finalmente, la agregación del sentimiento a nivel de sesión y el análisis de su dinámica interna contribuyen a una caracterización más robusta del tono emocional del discurso, reduciendo sesgos derivados de la longitud de las intervenciones y fortaleciendo la interpretabilidad de los resultados.

Pese a los resultados obtenidos, este estudio presenta una serie de limitaciones que deben ser consideradas al interpretar sus conclusiones. En primer lugar, el enfoque temático se restringe a cinco ejes sustantivos, lo que implica que ciertos contenidos relevantes del debate parlamentario, como materias político- institucionales, constitucionales o procedimentales, quedan fuera del análisis al ser clasificados como “Otros”.

En segundo lugar, el análisis de sentimiento se basa en una clasificación de polaridad (negativo, neutral y positivo), lo que bien resulta adecuada para una caracterización

general, no permite capturar matices emocionales más específicos, como indignación, preocupación, ironía o esperanza. Además, aún cuando BETO fine-tuned mostró un desempeño elevado, los modelos de análisis de sentimiento presentan limitaciones inherentes al procesamiento automático de lenguaje, especialmente en contexto donde la evaluación crítica se expresa mediante registros técnico, jurídicos o institucionales que pueden diluir marcadores emocionales más explícitos.

Asimismo, el análisis se realiza a nivel agregado de sesiones y bloques, sin considerar diferencias entre actores individuales, partidos políticos o roles institucionales, lo que podría ocultar variaciones relevantes en estilos discursivos específicos.

Finalmente, aunque el enfoque léxico-demostrativo es robusto y transparente, depende de la cobertura y calidad del vocabulario definido, lo que implica que ciertos conceptos emergentes o cambios semánticos podrían no ser capturados plenamente.

A partir de las limitaciones identificadas, se abren múltiples líneas de investigación futura. En primer lugar, el diccionario temático podría ampliarse para incorporar nuevos ejes del debate parlamentario, como materias constitucionales, político-institucionales o de derechos humanos, permitiendo una cobertura más exhaustiva del discurso legislativo. En segundo lugar, el análisis de sentimiento podría enriquecerse mediante la incorporación de modelos orientados a la detección de emociones específicas, o mediante enfoques supervisados entrenados directamente sobre corpus legislativos chilenos, lo que permitiría capturar con mayor precisión la complejidad emocional del debate parlamentario.

Asimismo, futuras investigaciones podrían explorar el cruce entre tópico, sentimiento y actor político, analizando diferencias discursivas entre partidos políticos, coaliciones o periodos legislativos. Finalmente, la aplicación del enfoque desarrollado en este trabajo a otros cuerpos legislativos, como la Cámara de Diputados o parlamentos comparados, permitiría evaluar la generalización del modelo y avanzar hacia análisis comparativos del discurso político institucional.

En síntesis, este estudio demuestra que el análisis automatizado del discurso parlamentario, cuando se sustenta en una metodología rigurosa y validada, permite revelar patrones temáticos y emocionales relevantes para la comprensión del funcionamiento político-institucional, contribuyendo tanto al campo de análisis político como al desarrollo aplicado de técnicas de procesamiento de lenguaje natural en contextos especializados.

Bibliografía

- Núñez, C. (2012). Descortesía verbal y conflicto político: El caso del discurso parlamentario chileno. *RLA. Revista de Lingüística Teórica y Aplicada*, 50(1), 131–156.
- De Cock, B. (2012). Cortesía, descortesía y discurso político: Un análisis contrastivo del discurso parlamentario español y chileno. *RLA. Revista de Lingüística Teórica y Aplicada*, 50(2), 89–110.
- Horn, C., & Balenzuela, P. (2025). Multilingual NLP for political text analysis in Latin America: Methods, challenges and applications.
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J., Kang, H., & Rivas, J. (2020). Spanish pre-trained BERT model and evaluation data. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Bravo-Marquez, F., Pérez, J., Díaz, G., Aponte, X., & Aranda, P. (2023). Transformers for Spanish sentiment analysis: A benchmark and new fine-tuned models.
- Pérez, J. M., Pérez, J., & Rosso, P. (2021). pysentimiento: A python toolkit for SocialNLP tasks. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 134–142.