



Análisis de Sentimientos sobre Titulares de Noticias

Alfredo Salvador Bistrain Montiel

Introducción a la Ciencia de Datos
Proyecto Final

1 Introducción

La dinámica informativa contemporánea está marcada por una producción continua de titulares que buscan captar la atención del lector. Estos encabezados no solo comunican hechos: condensan emociones, enfoques narrativos y, de manera más sutil, sesgos editoriales que pueden moldear percepciones colectivas. Comprender cómo diferentes medios construyen emocionalmente sus titulares constituye un reto estadístico y computacional significativo, especialmente cuando se carece de datos etiquetados y se trabaja con texto breve, ruidoso y altamente variable.

En este proyecto, se propone un planteamiento para medir el perfil emocional y narrativo de los medios de comunicación mexicanos. Donde se integran técnicas modernas de representación lingüística y modelos neuronales interpretables. Combinando tres pilares metodológicos:

1. **Análisis léxico-emocional** basado en el NRC Emotion Lexicon para cuantificar señales afectivas (*valence* y *arousal*) presentes en miles de titulares.
2. **Reducciones no lineales de dimensión** (PCA) que revelan estructuras latentes en el espacio emocional de los medios.
3. **Un modelo LSTM enriquecido con *embeddings* por dominio**, capaz de aprender simultáneamente patrones lingüísticos y representaciones vectoriales propias de cada medio, permitiendo recuperar para cada uno un conjunto de parámetros interpretables que actúan como un “perfil emocional” o β_i editorial refiriéndonos con β_i a un concepto similar a lo que se hace en regresión lineal simple. Más sobre esto en la sección correspondiente.

Este enfoque que combina *embeddings* estructurales, emocionales y contextuales permite no solo clasificar titulares, sino también descomponer el efecto del lenguaje versus el efecto del medio, abriendo la puerta a una lectura más transparente e interpretativa de modelos neuronales usualmente considerados “cajas negras”.

Los resultados obtenido muestran que los medios presentan huellas emocionales distintivas, recuperables a partir de sus titulares incluso sin etiquetas manuales. Además, los *embeddings* por dominio revelan agrupamientos coherentes entre medios y diferencias sistemáticas en su propensión a enfatizar tonos negativos, positivos o neutros. Este análisis, que se apoya únicamente en datos públicos, proporciona una metodología robusta, reproducible y extensible para el estudio cuantitativo del discurso mediático.

2 Marco teórico

2.1 Modelos lexicográficos de emociones

Los métodos lexicográficos constituyen una de las aproximaciones más directas para cuantificar la carga afectiva del lenguaje. En este enfoque, cada palabra del vocabulario se vincula con una o varias categorías emocionales predefinidas, generando representaciones basadas en conteos de ocurrencias.

Entre los recursos más influyentes destaca el *NRC Emotion Lexicon*, desarrollado por Mohammad and Turney (2013), el cual asigna etiquetas binarias a cada término según diez categorías emocionales: **anger**, **fear**, **joy**, **sadness**, **disgust**, **surprise**, **anticipation**, **trust**, así como las dimensiones **positive** y **negative**. La disponibilidad del lexicón en español permite su aplicación directa al análisis afectivo de medios latinoamericanos sin necesidad de modelos entrenados en inglés.

Construcción de métricas para evaluación de sentimientos: valence y arousal

A partir de los vectores de conteos proporcionados por el lexicón, es posible derivar métricas inspiradas en los modelos psicológicos dimensionales de emoción. En particular, el sistema PAD (Pleasure–Arousal–Dominance) postula que la experiencia afectiva puede representarse mediante tres ejes continuos, entre los cuales la dimensión de *pleasure* (valencia) y la de *arousal* (activación) son especialmente relevantes para el análisis textual. Este estudio adopta las siguientes definiciones:

- **Valence** (positividad–negatividad): mide el balance entre emociones asociadas a placer, confianza o expectativa positiva, frente a emociones de amenaza, aversión o malestar. Se calcula como:

$$\text{valence} = (\text{joy} + \text{trust} + \text{anticipation} + \text{positive}) - (\text{anger} + \text{fear} + \text{sadness} + \text{disgust} + \text{negative}).$$

Esta formulación refleja la polaridad emocional del titular y se alinea con la idea, ampliamente utilizada, de que la valencia puede representarse como la diferencia entre afecto positivo y negativo.

- **Arousal** (nivel de activación): estima cuán intensa, urgente o movilizadora es la carga emocional. Emociones como **anger**, **fear** y **surprise** se consideran de alta activación en la literatura psicológica, mientras que **anticipation** ha sido asociada a estados de alerta y preparación. Por ello, se define:

$$\text{arousal} = \text{anger} + \text{fear} + \text{surprise} + \text{anticipation}.$$

Esta métrica captura la tensión y dramatismo característicos del estilo noticioso.

2.2 Representaciones de texto y reducción de dimensión

Una vez que cada titular se convierte en un vector emocional, es posible estudiar la estructura del espacio afectivo mediante técnicas de reducción de dimensión. Estos métodos permiten detectar patrones globales, clústeres y relaciones entre medios que no son visibles directamente en el espacio original de diez dimensiones emocionales.

PCA: estructura lineal del espacio emocional

El *Principal Component Analysis* (PCA) constituye una herramienta fundamental por su capacidad de producir representaciones lineales interpretables. El objetivo central de PCA es encontrar combinaciones lineales de las variables originales que expliquen la mayor parte de la varianza. En el contexto de emociones:

- El **primer componente principal** suele representar un eje de polaridad emocional (positivo vs. negativo), pues captura las correlaciones naturales entre **joy/trust** y **anger/fear/sadness**.
- El **segundo componente principal** tiende a asociarse con emociones de activación, separando titulares más intensos o urgentes de aquellos más neutros o informativos.

PCA permite, por tanto, identificar de forma directa y explicable los ejes latentes más influyentes en el discurso noticioso. Esta propiedad lo convierte en el método principal de exploración estructural en este proyecto.

2.3 Modelos neuronales para secuencias

Las redes neuronales recurrentes, en particular las *Long Short-Term Memory* (LSTM), se han consolidado como herramientas eficaces para modelar secuencias lingüísticas. Una LSTM procesa el texto palabra por palabra, manteniendo un estado interno capaz de capturar dependencias de largo plazo. Su variante bidireccional, empleada en este estudio, incorpora tanto el contexto previo como el posterior, lo cual resulta especialmente adecuado para titulares breves y altamente informativos.

El modelo utilizado extiende la arquitectura tradicional mediante la incorporación de **embeddings de dominio**. Esta técnica, inspirada en enfoques de *multi-task learning* y *conditional embeddings*, consiste en aprender un vector latente para cada medio (e.g., [milenio.com](#), [proceso.com.mx](#), etc.). Dicho vector actúa como un *sesgo aprendido* (Kim et al. (2016)) que captura características propias del estilo informativo del medio: tono, nivel de emocionalidad, estilo narrativo, orientación general o tendencia editorial.

2.4 Interpretabilidad de modelos neuronales

Uno de los principales problemas de los modelos neuronales al ser adoptados es que, por su naturaleza, se hace un intercambio significativo entre poder de predicción e interpretabilidad. Sin embargo, al incluir factores explícitos como los embeddings de dominio, la arquitectura adquiere un grado adicional de interpretabilidad. El análisis de dichos embeddings permite estudiar si ciertos dominios tienden a amplificar emociones negativas o positivas. Además de tendencias estructurales en la forma de narrar noticias, independientes del contenido del titular.

Esta capacidad de interpretar parámetros específicos del modelo complementa los análisis lexicográficos y de reducción de dimensión, y contribuye a una comprensión más completa del ecosistema informativo.

3 Metodología

La metodología seguida en este estudio integra se puede resumir como: (i) la extracción sistemática de titulares de medios mexicanos, (ii) un proceso exhaustivo de limpieza y normalización de datos, y (iii) la construcción de representaciones emocionales y modelos neuronales capaces de capturar patrones narrativos. En esta sección se detalla cada etapa del pipeline.

3.1 Extracción de datos

3.1.1 Selección de fuentes

Uno de los objetivos centrales de este estudio es analizar las diferencias emocionales y estilísticas entre medios periodísticos con amplia presencia en el ecosistema informativo mexicano. Se adoptó una estrategia de filtrado por dominio. El conjunto final incluye diez medios digitales con cobertura nacional, trayectoria estable y reputación consolidada en los géneros de política, investigación, análisis económico y actualidad general. Estos son

- animalpolitico.com
- aristeguinoticias.com
- eleconomista.com.mx
- elfinanciero.com.mx
- elimparcial.com
- eluniversal.com.mx
- jornada.com.mx
- milenio.com
- proceso.com.mx
- excelsior.com.mx

3.1.2 Consulta al API de GDELT

La extracción de titulares se realizó mediante la API de GDELT (Global Database of Events, Language and Tone), la cual permite recuperar artículos en ventanas temporales específicas. Para cada día, el procedimiento produjo un conjunto independiente de titulares. Se almacenaron columnas:

- **date** (fecha YYYY-MM-DD)
- **title** (texto original)
- **domain** (fuente estándar en minúsculas)
- **url**

3.2 Limpieza y normalización de datos

3.2.1 Normalización del texto

Se aplicaron transformaciones a cada titular:

1. Conversión a minúsculas.
2. Eliminación de marcas editoriales (“| Opinión”, “– Video”, etc.).
3. Eliminación de espacios redundantes.
4. Sustitución de caracteres extraños.

3.2.2 Eliminación de duplicados

Esto evita que múltiples versiones de una misma nota alteren las métricas emocionales.

3.2.3 Filtrado por dominio válido

GDELT puede capturar piezas que provienen de versiones móviles, sindicadas o espejos de la misma nota. Para garantizar consistencia, se mantuvieron únicamente dominios pertenecientes explícitamente a la lista seleccionada.

3.2.4 Balanceo por fuente

La distribución original de titulares por medio presentaba una fuerte desigualdad: algunos periódicos tenían registros en la API de entre 40 000 y 60 000 notas al año, mientras que otros menos de 10 000. En particular, se identificó el medio con menor número de titulares válidos después de la limpieza (`animalpolitico`, con aproximadamente 9400 notas), y se utilizó este valor como límite uniforme. Este proceso garantiza que:

1. ninguna fuente domina el entrenamiento de la red neuronal;
2. los embeddings de dominio β_i se aprenden a partir de una cantidad comparable de ejemplos por medio;
3. la comparación posterior entre estilos editoriales es justa y estadísticamente interpretable.

Además, el balanceo evita que los modelos de análisis de dimensión (PCA) agrupen medios simplemente por volumen, permitiendo que las estructuras encontradas reflejen diferencias editoriales reales y no únicamente la densidad de datos.

3.3 Preparación del dataset para la red neuronal

3.3.1 Tokenización

Para este proyecto se empleó una tokenización deliberadamente simple, diseñada para procesar titulares periodísticos sin introducir transformaciones agresivas. Cada texto se convierte a minúsculas y se extraen únicamente secuencias alfábéticas, incluyendo caracteres propios del español (tildes y diéresis). Esta operación se implementa mediante una expresión regular que retiene únicamente palabras válidas:

$$\text{tokens} = \text{re.findall}([a-zA-Z]+).$$

Este esquema tiene tres ventajas principales: (1) elimina puntuación, enlaces y símbolos que no aportan carga afectiva; (2) preserva adecuadamente la morfología del español; (3) genera un vocabulario compacto, estable y adecuado para ser usado por modelos basados en secuencias como LSTM.

3.3.2 Asignación de etiquetas emocionales

Para entrenar el modelo de clasificación fue necesario convertir métrica de *valence* en una etiqueta discreta. Con este propósito se definió una regla de decisión basada en umbrales simétricos:

$$y = \begin{cases} 0, & \text{si } \text{valence} < -\tau, \\ 1, & \text{si } -\tau \leq \text{valence} \leq \tau, \\ 2, & \text{si } \text{valence} > \tau, \end{cases} \quad y \in \{0, 1, 2\}.$$

Este procedimiento refleja la intuición de que valores muy negativos de *valence* corresponden a emociones aversivas (e.g. `anger`, `fear`, `sadness`), mientras que valores muy positivos sugieren emociones de apertura o confianza (e.g. `joy`, `trust`). Los titulares cuya polaridad se mantiene cerca de cero se consideran neutros.

En este estudio se utilizó un umbral fijo $\tau = 1$, el cual deja que la propia señal lexicográfica determine el signo de la emoción percibida. No obstante, la regla puede generalizarse a otros umbrales aprendidos o determinados empíricamente.

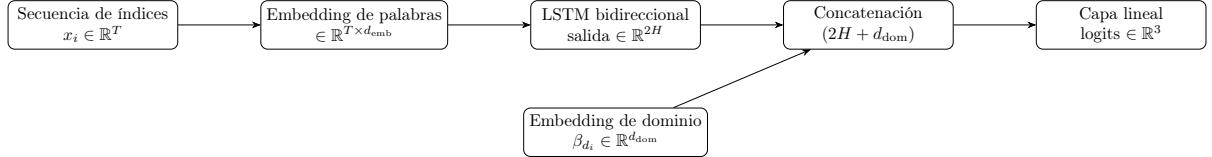
3.4 Entrenamiento del modelo LSTM

El modelo final corresponde a una arquitectura LSTM bidireccional con dos fuentes de información: (i) el contenido lingüístico del titular, codificado como una secuencia de índices; y (ii) un embedding aprendido para cada medio de comunicación, que capta el estilo editorial y posibles sesgos sistemáticos. De manera paralela, el identificador entero de la fuente se transforma mediante una capa de *embedding* en un vector denso:

$$d_i \longrightarrow \beta_{d_i} \in \mathbb{R}^{d_{\text{dom}}},$$

donde d_i es el identificador entero del medio al que pertenece cada titular; $\beta_{d_i} \in \mathbb{R}^{d_{\text{dom}}}$ es el vector latente aprendido que resume el estilo editorial característico de ese medio; H denota el tamaño del estado

oculto de cada dirección de la BiLSTM, por lo que la representación textual obtenida del titular es $h_{\text{text}} \in \mathbb{R}^{2H}$; y la concatenación $h_{\text{final}} = [h_{\text{text}} \parallel \beta_{d_i}] \in \mathbb{R}^{2H+d_{\text{dom}}}$ combina simultáneamente la información semántica del texto y el sesgo editorial capturado por el embedding de dominio. Esta combinación permite capturar interacciones entre lo que se dice en el titular y la forma típica en la que cada medio enmarca la información. El procesamiento que realiza la red puede resumirse de la siguiente manera:



El objetivo es que la LSTM capture la información semántica del titular mientras que el embedding de dominio introduce un sesgo aprendido que refleja la línea editorial o estilo emocional característico de cada medio. La combinación de ambas señales permite que el modelo represente tanto el contenido del titular como el contexto institucional desde el cual se emite.

El modelo se entrenó durante 12 épocas utilizando el optimizador Adam con regularización L_2 , la pérdida de entropía cruzada y un *batch size* entre 32 y 64. Para garantizar estabilidad numérica se aplicó *gradient clipping* con norma máxima de 1.0 y, cuando fue posible, se aceleró el entrenamiento mediante CUDA. Este esquema permitió que la red aprendiera representaciones distribuidas del texto que capturan patrones sintácticos y semánticos relevantes, integrando de forma implícita la señal emocional proveniente del lexicón. Asimismo, los *embeddings* de dominio aprendieron factores latentes que reflejan el estilo editorial de cada medio y modulan los logits finales en combinación con la representación textual. En conjunto, el modelo no sólo clasifica titulares, sino que descompone la predicción en contenido emocional del texto y sesgos estilísticos propios de la fuente, proporcionando una caracterización interpretable del discurso mediático.

4 Resultados

4.1 Depuración de encabezados

La extracción inicial abarcó todos los titulares publicados por las diez fuentes seleccionadas durante el periodo comprendido entre enero de 2024 y noviembre de 2025. El proceso generó un total de 299,118 artículos. A partir de este volumen bruto, se aplicaron varias etapas de limpieza y filtrado con el objetivo de obtener un conjunto homogéneo, consistente y balanceado entre medios.

- Limpieza básica.** Se eliminaron entradas vacías, duplicados exactos, URLs inválidas y registros con títulos evidentemente defectuosos. Después de esta etapa se conservaron **288,959 artículos**.
- Revisión por dominio y depuración semántica.** Se filtraron los artículos para conservar únicamente aquellos provenientes de los dominios seleccionados como parte del análisis comparativo entre medios nacionales. Tras esta depuración, el total de encabezados quedó en **283,538 artículos**.
- Balanceo por fuente.** Con el fin de realizar comparaciones justas entre medios —sin que la capacidad predictiva del modelo fuera dominada por las fuentes con mayor volumen— se aplicó un muestreo estratificado uniforme. Para ello, se identificó la fuente con el menor número de titulares (animalpolitico.com, con 9,410 artículos), y se extrajo una muestra del mismo tamaño para cada uno de los diez medios.

Como resultado, se obtuvo un banco de palabras balanceado de:

$$10 \text{ medios} \times 9,400 \text{ titulares} = \mathbf{94,000 \text{ artículos}}$$

Este conjunto final constituye la base utilizada para el análisis emocional, las visualizaciones exploratorias y el entrenamiento del modelo LSTM con embeddings de dominio.

4.2 Análisis de Componentes Principales (PCA) del espacio emocional

Para visualizar la estructura latente de los titulares en términos de sus diez dimensiones emocionales NRC, se aplicó un Análisis de Componentes Principales (PCA). Los dos primeros componentes concentran una proporción considerable de la variabilidad total del espacio:

$$\text{Varianza explicada: } \text{PC1} = 0.3668, \quad \text{PC2} = 0.2805.$$

En conjunto, estos dos ejes capturan más del **64%** de la estructura emocional del banco de palabras, lo que permite una visualización bidimensional informativa.

Interpretación del Componente Principal 1 (PC1)

El primer componente refleja un eje claro de **polaridad emocional**. Valores altos de PC1 están asociados con titulares con carga afectiva positiva, mientras que valores bajos corresponden a narrativas fuertemente negativas. Esto puede observarse al inspeccionar ejemplos extremos:

- **Titulares negativos de alta activación**, como:
“imputan a donovan n, segundo detenido por asesinato del abogado david cohen; homicidio calificado; queda en prisión preventiva”
 presentan una valencia muy baja (valence = -19) y alta activación (arousal = 12), ubicándose en la región más negativa del eje PC1.
- **Titulares positivos**, por ejemplo:
“trump busca acuerdo de paz con irán tras ataques conjuntos con israel; netanyahu promete lograr la paz juntos”
 exhiben una valencia alta (valence = 21) y activación moderada (arousal = 8), situándose hacia los valores elevados del componente.

Interpretación del Componente Principal 2 (PC2)

El segundo componente captura variabilidad independiente relacionada principalmente con el **nivel de activación emocional (arousal)**. Titulares que combinan emociones intensas como **anger**, **fear** o **surprise** tienden a aparecer con valores altos de PC2, aun cuando su valencia sea negativa o positiva. Por el contrario, titulares descriptivos o administrativos se concentran en valores bajos o moderados.

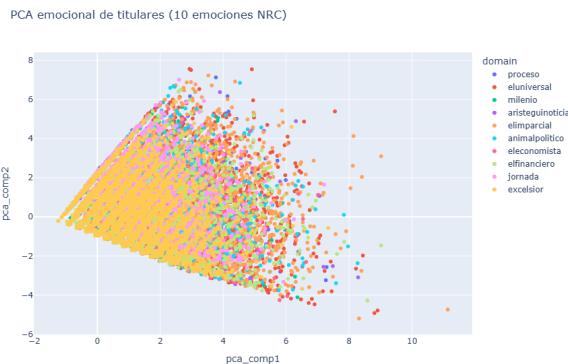


Figure 1: Proyección PCA de los titulares sobre el espacio emocional del NRC (10 dimensiones). Cada punto representa un titular, coloreado por su medio de origen.

Estos patrones justifican la construcción de un clasificador supervisado: existe una estructura separable (aunque no linealmente perfecta) que un modelo puede aprovechar para predecir la carga afectiva de cada titular. A partir de esta observación, procedemos a entrenar un modelo LSTM bidireccional mejorado con embeddings de dominio, con el fin de capturar simultáneamente: (i) el contenido lingüístico, (ii) el patrón emocional latente y (iii) el estilo característico de cada medio.

4.3 Resultados del modelo LSTM con embeddings de dominio

El modelo se entrenó durante 12 épocas utilizando el conjunto balanceado de 94 000 titulares (9 400 por medio).

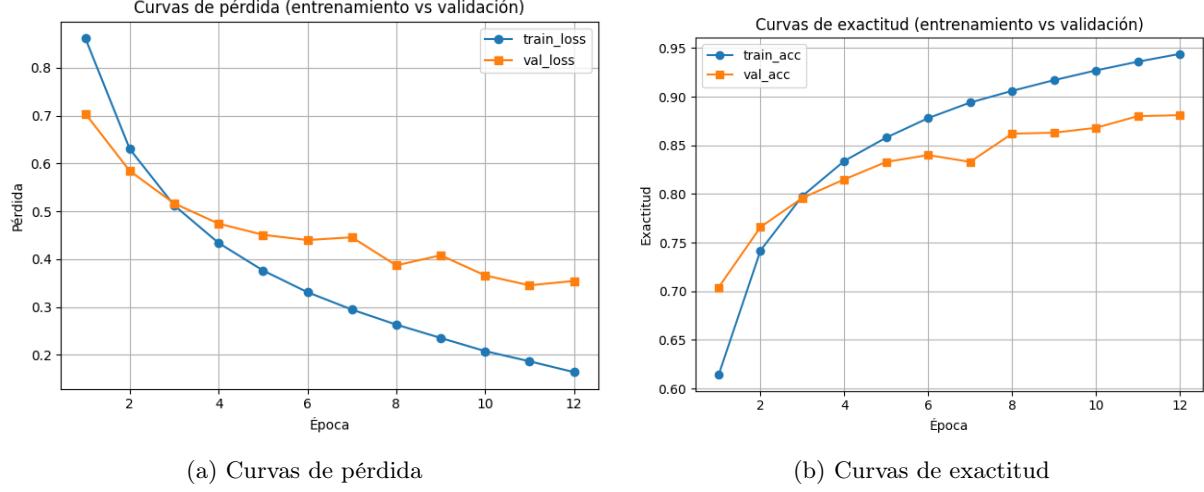


Figure 2: Evolución del entrenamiento: pérdida (izquierda) y exactitud (derecha).

El comportamiento observado es estable y sin señales de sobreajuste severo: la pérdida desciende de forma monótona y la exactitud de validación aumenta gradualmente hasta estabilizarse alrededor de 0.88. En términos de desempeño, el clasificador alcanza una exactitud del **88.1%** en validación, lo cual es notable considerando que las etiquetas emocionales provienen de un *proxy lexicográfico* y no de anotaciones humanas. El modelo aprende patrones emocionales robustos pese al ruido inevitable del método de etiquetado.

Este comportamiento también valida el uso de embeddings de dominio: el modelo logra integrar el estilo editorial de cada medio dentro de su predicción afectiva, sin que esto derive en sobreajuste. Más adelante se presentan los parámetros aprendidos β_i para cada fuente y su interpretación.

Interpretación de la matriz de confusión

La matriz obtenida es:

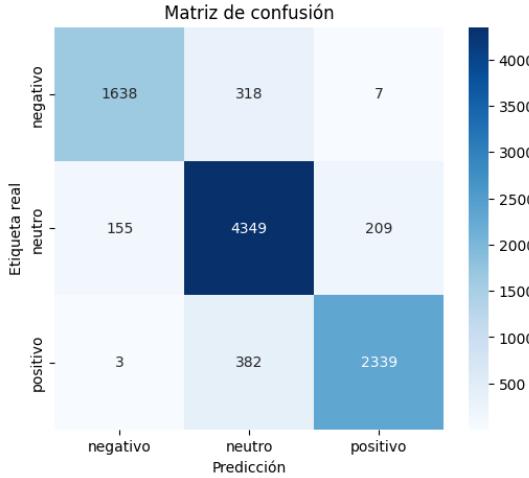


Figure 3: Matriz de confusión del modelo LSTM con embeddings de dominio. La diagonal indica los aciertos por clase y las celdas fuera de la diagonal corresponden a errores de clasificación.

Clase negativa. El modelo clasifica correctamente 1638 titulares negativos. Los errores ocurren casi exclusivamente hacia la clase neutra (318 casos), mientras que solo 7 se confunden como positivos. Esto

indica que el modelo distingue bien el lenguaje negativo, y cuando se equivoca, tiende a suavizar la predicción hacia la neutralidad.

Clase neutra. Es la clase más numerosa y también la mejor clasificada: 4349 aciertos. Los errores se dividen en negativos (155) y positivos (209). Esto sugiere que la red neuronal aprendió de manera efectiva patrones lingüísticos que caracterizan titulares neutrales dentro del periodismo mexicano.

Clase positiva. El modelo acierta 2339 titulares positivos, pero confunde 382 como neutros. Solo 3 se clasifican erróneamente como negativos. La frontera entre lo positivo y lo neutro es la más difusa, algo esperable dado que muchos titulares “positivos” contienen lenguaje informativo o técnico sin entusiasmo explícito.

De aquí podemos notar ciertos patrones

- El modelo es conservador ante la incertidumbre, tiende a predecir la clase neutra.
- Las clases positiva y negativa casi nunca se confunden entre sí, lo cual indica que la red aprendió una separación robusta entre las polaridades extremas.
- La mayor parte de los errores está en la frontera neutro–positivo, consistente con el estilo editorial de los titulares.
- El desempeño global muestra un modelo estable, interpretable y adecuado para el análisis de narrativa mediática.

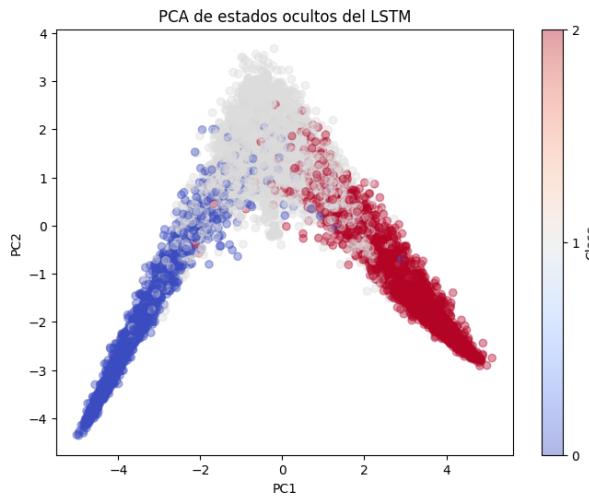


Figure 4: Proyección PCA de los estados ocultos finales del LSTM. Cada punto corresponde al vector oculto asociado a un titular, coloreado según su etiqueta emocional.

Interpretación del PCA de estados ocultos

El PCA aplicado a los estados ocultos revela que el LSTM aprendió una organización latente altamente estructurada del espacio emocional. Los titulares negativos y positivos se separan en dos brazos opuestos, lo cual indica que la red reconstruyó un eje interno análogo a la *valencia emocional*. En contraste, los titulares neutros forman un cúmulo compacto en la zona central, reflejando su menor carga afectiva y un estilo lingüístico más plano. La geometría resultante —en forma de “V”— sugiere que el modelo distingue un continuo emocional donde los extremos corresponden a emociones intensas, mientras que el vértice agrupa contenido descriptivo o administrativo.

Esta separación clara confirma que la red no solo clasifica, sino que sintetiza relaciones semánticas profundas: titulares con patrones similares terminan en regiones próximas del espacio oculto, aun si el etiquetado proviene de un proxy lexicográfico. La estructura obtenida muestra que el modelo aprendió representaciones internas coherentes con modelos psicológicos como PAD (Pleasure–Arousal–Dominance), pese a no haber sido entrenado explícitamente para ello.

¿Por qué aplicar PCA a los estados ocultos?

El vector oculto final del LSTM, $h_i \in \mathbb{R}^{2H}$, condensa la información que el modelo considera relevante del titular. Aplicar PCA permite visualizar cómo la red organiza este espacio latente y evaluar si las clases se separan antes de la capa final. Este análisis cumple tres funciones principales: (i) verificar que el modelo aprendió patrones emocionales consistentes; (ii) identificar un eje dominante asociado a la polaridad negativa–positiva; y (iii) evaluar si los titulares neutros ocupan una región estable y diferenciada. La estructura observada confirma que el aprendizaje fue sólido y que el espacio oculto refleja una geometría emocional interpretable.

4.4 Efecto específico del medio: los vectores β_i

Los embeddings de dominio β_i representan el estilo editorial característico de cada medio y actúan como un sesgo aprendido que se suma a los logits finales del modelo. De este modo, cada vector

$$\beta_i \in \mathbb{R}^3$$

modifica la probabilidad de predecir las clases negativa, neutra o positiva, independientemente del contenido del titular. Sea $B \in \mathbb{R}^{N_{\text{dom}} \times d_{\text{dom}}}$ la matriz de embeddings de dominio aprendida por el modelo. Cada medio i posee un identificador entero d_i , y su vector latente se obtiene como la d_i -ésima fila de B :

$$\beta_{d_i} = B[d_i, :] = e_{d_i}^\top B,$$

donde e_{d_i} es el vector estándar (one-hot) con un 1 en la posición correspondiente al dominio. Este embedding se concatena con la representación textual del LSTM:

$$h_{\text{final}} = \begin{bmatrix} h_{\text{text}} \\ \beta_{d_i} \end{bmatrix} \in \mathbb{R}^{2H+d_{\text{dom}}}.$$

Si la matriz de pesos de la capa lineal se descompone como $W = [W_{\text{text}} \mid W_{\text{dom}}]$, con $W_{\text{text}} \in \mathbb{R}^{3 \times 2H}$ y $W_{\text{dom}} \in \mathbb{R}^{3 \times d_{\text{dom}}}$, los logits finales se obtienen mediante:

$$\text{logits} = W_{\text{text}} h_{\text{text}} + W_{\text{dom}} \beta_{d_i} + b.$$

Por lo tanto, las cantidades reportadas como β_i corresponden directamente a las filas de la matriz de embeddings B , que es un parámetro entrenable optimizado por *backpropagation*. Estas filas representan el efecto específico de cada medio sobre los logits finales del clasificador. En el contexto de redes neuronales, se les denomina *logits* a las salidas crudas de la última capa lineal del modelo, es decir, los puntajes reales que aún no han sido transformados en probabilidades. Estos valores pueden ser positivos o negativos y representan la evidencia relativa a favor de cada clase antes de aplicar la función **softmax**. Posteriormente, el **softmax** normaliza los logits para generar una distribución de probabilidad válida. En nuestro modelo, los vectores β_i asociados a cada medio se suman directamente sobre los logits, actuando como un desplazamiento (*bias*) específico del dominio que modifica los puntajes de cada clase antes de la normalización final.

Los patrones de la Figura 5 muestran que todos los medios tienen valores negativos en las dimensiones **neg** y **pos**, lo que indica que ninguno empuja por sí mismo hacia emociones extremas: la carga afectiva proviene principalmente del texto. En contraste, varios medios presentan valores positivos en la dimensión neutral —especialmente **eluniversal**, **proceso** y **jornada**— lo cual sugiere un estilo más informativo y menos emocionalizado que incrementa la probabilidad de clasificar un titular como neutro cuando su contenido no es claramente positivo o negativo.

Medios como **animalpolitico** o **excelsior** muestran β_{neu} más bajos, indicando que su forma de escribir titulares deja más peso al contenido lingüístico para determinar la emoción. En conjunto, los vectores β_i no reflejan ideología ni sesgo político, sino diferencias sistemáticas en neutralidad, tono y estructura narrativa. Esto permite descomponer la decisión del modelo en:

emociones del texto + estilo propio del medio,

proporcionando una interpretación cuantitativa del estilo editorial en la prensa mexicana.

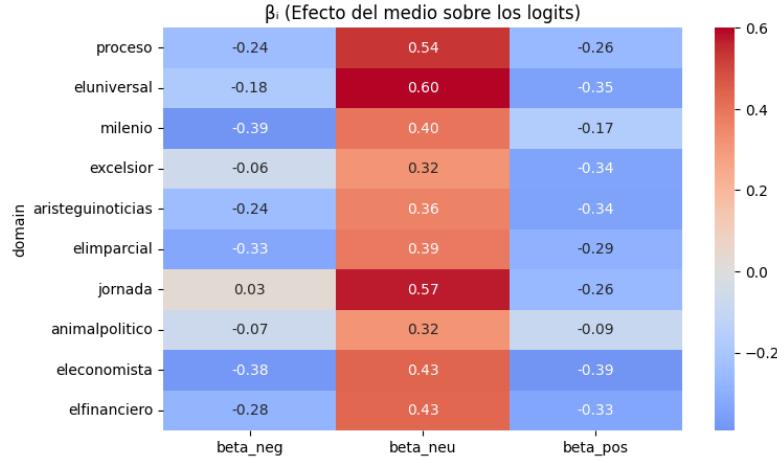


Figure 5: Efectos específicos de cada medio sobre los logits finales (β_i) para las clases negativa, neutra y positiva.

5 Discusión y conclusiones

El análisis combinado basado en un lexicón emocional, reducción de dimensión y un modelo LSTM con embeddings de dominio permitió caracterizar de forma cuantitativa la carga afectiva de los titulares y las diferencias estilísticas entre medios nacionales. Los resultados muestran que la mayor parte de los encabezados se concentra alrededor de valores moderados de *valence* y *arousal*, lo cual es consistente con el carácter informativo de la prensa. Aun así, existen variaciones pequeñas pero sistemáticas entre medios, sugiriendo estilos editoriales diferenciados.

El PCA aplicado a las emociones lexicográficas mostró una estructura difusa y parcialmente mezclada entre clases, mientras que el PCA de los estados ocultos del LSTM reveló una separación mucho más clara entre polaridades negativas, neutras y positivas. Esto indica que el modelo aprendió una representación interna más rica que integra patrones sintácticos y semánticos no capturados por conteos lexicográficos, organizando los titulares a lo largo de un eje emocional similar al de la teoría psicológica de valencia.

Los embeddings de dominio proporcionaron un componente interpretativo adicional: todos los medios mostraron efectos negativos moderados en las clases extremas y valores positivos en la clase neutra, lo cual sugiere que, en ausencia de señales emocionales fuertes, el estilo editorial tiende a favorecer interpretaciones neutrales. Medios como *eluniversal*, *proceso* o *jornada* muestran mayor inclinación hacia neutralidad estilística, mientras que otros presentan un efecto menor, dejando el peso de la predicción casi por completo en el contenido del titular.

En conjunto, estos resultados demuestran que la combinación de métodos lexicográficos, PCA y modelos neuronales permite obtener una visión coherente y complementaria del discurso afectivo en medios nacionales. La carga emocional explícita en los titulares es moderada, pero existen patrones editoriales detectables que influyen en la probabilidad de clasificar un texto como negativo, neutro o positivo. Este enfoque abre la puerta a extensiones futuras como modelos contextuales (BERT), análisis temporal y adaptación por dominio, que permitirían profundizar en el estudio emocional y narrativo del periodismo mexicano.

References

- Daumé III, H. (2007). Frustratingly easy domain adaptation. *ACL*.
- Kim, Y., Jernite, Y., Sontag, D., & Rush, A. (2016). Domain-sensitive hierarchical neural networks for document modeling. *ACL*.
- Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(4), 261–292.
- Mohammad, S. M., & Turney, P. D. (2013). Nrc emotion lexicon [Accessed: 2025-02-19]. *National Research Council, Canada*. <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- Plank, B., & Søgaard, A. (2019). Embedding domain knowledge via domain tags for neural nlp. *ACL*.