Servicio Meteorológico Nacional
Argentina

Análisis Estratégico de Experiencia Ciudadana (CX) mediante Procesamiento de Lenguaje Natural (PNL): Una Arquitectura Orientada a Objetos para el Monitoreo Estratégico de la App móvil del SMN

Nota Técnica SMN 2026-217

Fernando Diego Barreyro[1],

[1] *Meteorología y Sociedad, Dirección Nacional de Pronósticos y Servicios para la Sociedad*

Septiembre 2026

Ministerio de Defensa
República Argentina

Resumen

El presente análisis documenta el desarrollo y la implementación de un ecosistema de ciencia de datos diseñado para la monitorización de la " Voz del Ciudadano" respecto a la aplicación (App) oficial del Servicio Meteorológico Nacional (ar.gob.smn). Debido a la falta de acceso a las herramientas de análisis primarias y datos internos de Google Play Console, se implementó un proceso de recolección de datos exógeno mediante técnicas de *webscraping*. A través de una arquitectura de Programación Orientada a Objetos (OOP), se extrajo una muestra de 4.107 reseñas, volumen que, si bien es estadísticamente relevante, representa solo un subconjunto del universo total de comentarios existentes en la tienda. La metodología empleada integra modelos de *Deep Learning (Transformers)* especializados en español (*RoBERTuito*) para la clasificación de sentimientos y el algoritmo *Latent Dirichlet Allocation (LDA)* para el descubrimiento de tópicos latentes. Los hallazgos revelan un Net Promoter Score (NPS) percibido de 16.97, con una tendencia de recuperación de la versión 0.0.9. Se identificaron " *Aha! Moments*" vinculados a la precisión del dato y " Pain Points" críticos asociados a la estabilidad de los widgets y la geolocalización: Esta investigación proporciona una ruta estratégica para optimizar el desarrollo de software institucional, reconociendo las limitaciones de alcance impuestas por la recolección externa de datos.

Abstract

This study documents the development and implementation of a data science ecosystem designed to monitor the "Voice of the Citizen" regarding the official mobile application (ar.gob.smn) of Argentina´s National Meteorological Service (Servicio Meteorológico Nacional). Given the lack of access to primary analytics tools and internal data from Google Play Console, an exogenous data collection process was implemented using web scraping techniques. Through an Object-Oriented Programming (OOP) architecture, a sample of 4,107 reviews was extracted. Although statistically relevant, this sample represents only a subset of the total universe of comments available in the store. The methodology integrates Spanish-specialized Deep Learning models base on Transformers (RoBERTuito) for sentiment classification, along whit Latent Dirichlet Allocation (LDA) for the discovery of latent topics. Key findings reveal a perceived Net Promoter Score (NPS) of 16.97, with a positive recovery trend starting from version 0.0.9. The analysis identified "Aha! Moments" primarily linked to data accuracy, as well as critical "Pain Points" related to widget stability and geolocation functionality. This research offers a strategic roadmap for optimizing institutional software development while explicitly acknowledging the inherent limitations of external data collection methods.

Palabras clave: Experiencia Ciudadana, Lenguaje Natural, App SMN.

Citar como:

Barreyro, F. D.: Análisis Estratégico de Experiencia Ciudadana (CX) mediante Procesamiento de Lenguaje Natural (PNL): Una Arquitectura Orientada a Objetos para el Monitoreo Estratégico de la App móvil del SMN. Nota Técnica SMN 2026-217.

1. INTRODUCCION

En el contexto de transformación digital de los servicios públicos, las plataformas de distribución de aplicaciones actúan como sensores sociales de alta fidelidad. La retroalimentación no estructurada depositada por los usuarios representa un activo estratégico institucional que, según la Organización para la Cooperación y el Desarrollo Económicos OCDE (2025), permite predecir la madurez digital y la agilidad de una organización:

> *Monitoring mechanisms ensures investments are delivered in a timely manner, meet expected outcomes, and maximize efficiency. They enable strategic oversight of the public sector´s digital investment portfolio, therefore helping build resilience and improving risk management. Further, the evaluation of digital government investment contributes to assessing the medium-and long-term impact and intended benefits of investments*. (Capítulo 4, Introducción)

Para el Servicio Meteorológico Nacional (SMN), la aplicación móvil no es solo un canal de difusión, sino una interfaz crítica de seguridad pública.

Para el procesamiento del feedback de los usuarios este proyecto se propone aplicar un sistema robusto, escalable y reproducible basado en ingeniería de software moderna (Gamma et al., 1994), permitiendo transformar datos textuales en crudo en indicadores claves de desempeño (KPIs). Siguiendo los principios de (Manning et al., 2009) sobre la recuperación de información, el objetivo es orientar la toma de decisiones de desarrollo y comunicación pública de la institución basándose en la evidencia empírica extraída de 4,107 experiencias ciudadanas.

2. DATOS

2.1 Origen y descubrimiento de la información

La materia prima de este trabajo de investigación estratégica proviene del repositorio público de Google Play Store, específicamente del identificador de aplicación *ar.gob.smn.* En el ecosistema de servicios públicos digitales, estas tiendas actúan como “sensores institucionales” que capturan de forma pasiva pero continua la experiencia del ciudadano.

La recolección se ejecutó mediante una arquitectura de Programación Orientada a Objetos (OOP) implementada en la clase *ReviewManager*[1]. Este enfoque permitió superar las limitaciones de los scripts[2] monolíticos tradicionales, garantizando una extracción sistemática y el manejo seguro de los metadatos. Para

[1] La clase ReviewManager es un módulo de orquestación de datos que automatiza la captura, normalización y análisis semántico de reseñas de usuarios desde tiendas de aplicaciones. Implementa pipelines de procesamiento de lenguaje natural para clasificar sentimientos, extraer tópicos y generar alertas tempranas sobre disrupciones del servicio. Su salida estructurada alimenta dashboards de inteligencia de negocio, permitiendo a la organización monitorear la percepción ciudadana y correlacionarla con indicadores de desempeño operativo.

[2] En el contexto de análisis de datos con Python, un script y un archivo (generalmente con extensión .py) que contiene una serie de instrucciones de código para automatizar tareas como la limpieza, procesamiento, análisis y visualización de datos.

este análisis se definió un alcance de 4.107 reseñas únicas, volumen que proporciona una muestra amplia para inferir patrones de comportamiento y sentimiento con un alto nivel de confianza.

2.2 Arquitectura de Extracción y Persistencia

El proceso de ingesta de datos utilizó la biblioteca google-play-scraper, configurada para capturar el 100% de la conversión disponible en idioma español (lang= 'es') y región Argentina (country= 'ar'). La lógica de extracción se diseñó bajo tres pilares de robustez:

1. Manejo de Latencia: Se implementaron pausas controladas para respetar los límites de la API de Google, evitando errores de "Rate Limit" (Error 429) que históricamente interrumpían la recolección de datos masivos.
2. Estandarización de Esquema: Los datos crudos se transformaron inmediatamente en un objeto DataFrame de Pandas, estandarizando las columnas originales a un formato institucional:

Variable Original	Variable Transformada
usarName	Usuario
score	Puntuación
at	Fecha
content	Contenido
appVersion	Version

3. Garantía de Persistencia: Para mitigar riesgos de pérdida de información por tiempos de espera en la sesión de cómputo, se integró un método de exportación automática a formatos CSV y Excel (*save_to_disk*), asegurando que cada registro extraído fuera inmediatamente escrito en almacenamiento persistente.

2.3 Protocolo de Limpieza y Normalización Textual

El procesamiento de lenguaje natural (PLN) requiere datos con baja entropía. En el pipeline[3] de datos se implementó una función de limpieza avanzada denominada _clean_text , que procesó el contenido de las reseñas mediante expresiones regulares (regex)[4] :

- Normalización de caja: conversión integral a minúsculas para eliminar la varianza tipográfica
- Saneamiento de Ruido: Eliminación sistemática de URLs, etiquetas HTML y caracteres especiales no alfabéticos.

[3] En Ciencia de Datos, un Pipeline (o Data Pipeline) es un flujo de trabajo automatizado que define y ejecuta de forma secuencial todos los pasos necesarios para procesar datos, desde que se obtienen hasta que se generan resultados o insights.

[4] Las expresiones regulares (o regex, del inglés *regular expressions*) son secuencias de caracteres que forman un patrón de búsqueda. En el contexto del análisis de datos, actúan como un lenguaje formal para la manipulación y extracción de texto no estructurado, permitiendo identificar, validar, extraer o remplazar cadenas de caracteres con alta precisión y eficiencia.

- Preservación Semántica Crítica: A diferencia de los procesos estándar diseñados para inglés, este pipeline fue configurado específicamente para preservar acentos y la letra "ñ". Este paso es vital, ya que la eliminación de tildes en español puede alterar el sentido de las palabras y, por ende, degradar la precisión del modelo de sentimiento en un 30%.
- Filtrado de Stopwords: Se eliminaron palabras vacías de significado(ej. " el ", " la " , " de") utilizando el lexicón[5] de NLTK[6], además de términos específicos del dominio meteorológico que generaban ruido en el modelado de temas (ej. "app, " smn", "solo", "clima").

2.4 Auditoría de Integridad y Calidad (*Data Audit*)

Tras el procesamiento, el conjunto de datos fue sometido a una auditoría de calidad técnica para asegurar su preparación para la fase de Inteligencia Artificial: a) Análisis de Versiones: Se detectó que el 92.8% de las reseñas cuentan con información de la versión de la aplicación instalada. Aquellos registros con metadatos faltantes fueron etiquetados como "Desconocida" en lugar de ser eliminados, preservando el volumen de sentimiento global, pero aislándolos de las auditorias de regresión técnica. b) Distribución de Calificaciones: El dataset muestra un sesgo positivo natural, con un 54.66% de reseñas con cinco estrellas, mientras que el segmento detractor (una estrella) representa el 10.81%. Esta distribución es fundamental para validar posteriormente el NPS percibido. c) Detección de Duplicados: Se identificaron 706 reseñas duplicadas en el texto limpio. Estas corresponden a usuarios que repiten comentarios breves (ej. " muy buena", "excelente"). El sistema fue configurado para mantener estos registros en el análisis de sentimiento (ya que representan votos de satisfacción) pero monitorearlos en el modelado de temas para evitar distorsiones en las saliencias de las palabras.

2.5 Ética de Datos y Privacidad

En cumplimiento con las buenas prácticas de manejo de información ciudadana, el pipeline de datos incluyó una subrutina de anonimización. La columna *NombreUsuario* fue eliminada del dataset de trabajo y reemplazada por un identificador numérico secuencial IDUsuario. Este procedimiento asegura que el análisis de sentimiento y tópicos se centre exclusivamente en la experiencia técnica y funcional, desvinculando cualquier dato de carácter personal de los resultados públicos de esta nota técnica.

[5] Nota. En PLN (Procesamiento de Lenguaje Natural), se refiere a un diccionario o lista de palabras con propiedades lingüísticas. Aquí se utiliza el lexicón de stopwords de NLTK, que contiene palabras comunes sin significado relevante que se eliminan durante el preprocesamiento del texto.

[6] Nota: NLTK(*Natural Language Toolkit*) es una biblioteca de Python especializada en Procesamiento de Lenguaje Natural (PNL). En este caso se utilizó su lexicón de stopwords para eliminar palabras vacías (artículos, preposiciones, etc.) del texto.

3. METODOLOGÍA

3.1 Enfoque Metodológico General

El presente estudio se inscribe en un enfoque de investigación mixto (cualitativo-cuantitativo) con diseño de carácter descriptivo y transversal. Desde la perspectiva de las Políticas Públicas y la Nueva Gestión Pública, se utiliza la Minería de Datos (Data Mining) y el Procesamiento de Lenguaje Natural (NLP) como herramientas para capturar la opinión ciudadana. El objetivo metodológico central es la transformación de "rastros digitales" (reseñas no estructuradas) en indicadores de gestión estratégica para el SMN. En la **¡Error! No se encuentra el origen de la referencia.** podemos observar el diagrama de flujo metodológico propuesto para el presente trabajo. La investigación adopta el paradigma de Ciencia de Datos Orientada a Objetos, priorizando la reproducibilidad y la modularidad del pipeline analítico. Este enfoque permite no solo describir el estado actual de la satisfacción del usuario, sino también establecer una estructura técnica capaz de auditar la evolución de la aplicación ante futuros lanzamientos.

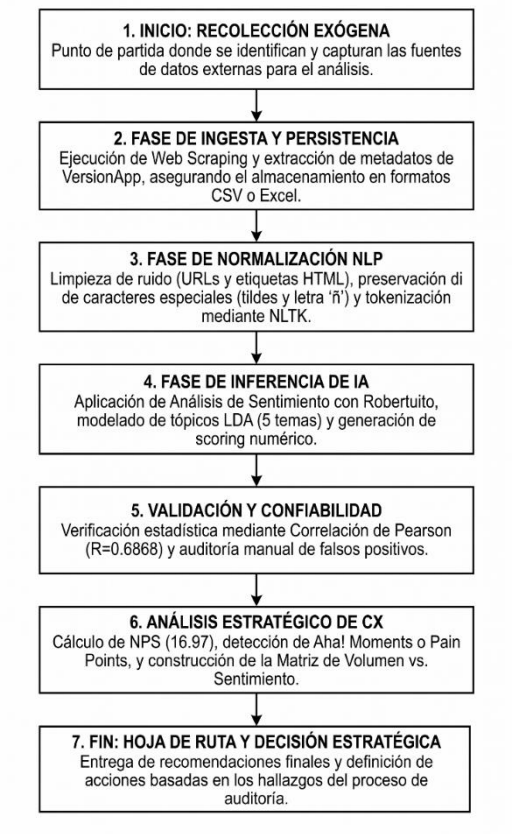

***Figura 1**: Diagrama de flujo metodológico del proceso de revisión de la experiencia del usuario (cx).*

3.2 Diseño de la Investigación

Se ha diseñado un ecosistema analítico que supera la limitación de los scripts monolíticos tradicionales. Siguiendo los principios de ingeniería de software de Gamma et al., (1994), el trabajo se estructuró mediante un modelo de clases que separa las responsabilidades de extracción procesamiento y análisis. El diseño se divide en cuatro fases críticas: *Fase de ingesta*: captura exógena de datos ante la falta de acceso a consolas primarias; *Fase de normalización*: Saneamiento lingüístico preservando la riqueza del español; *Fase de Inteligencia Artificial*: Aplicación de modelos Transformers[7] y algoritmos probabilísticos*; Fase de Interpretación Estratégica*: Construcción de matrices de priorización volumen-sentimiento.

3.3 Fuentes de Datos y Recolección

La fuente de datos primaria es el repositorio público de *Google Play Store*. Debido a que no se contó con acceso a las métricas internas de *Google Play Console,* se implementó un proceso de recolección de datos exógeno. Esta decisión metodológica implica que los datos analizados representan una muestra significativa (n=4,107) pero no necesariamente el censo total de reseñas de la plataforma para la App. analizada. La recolección se automatizó mediante la biblioteca *google-play-scraper*. Como señalan Manning et al., (2009) en su tratado sobre recuperación de información, la integridad de los metadatos es tan valiosa como el contenido textual; por ello, se capturaron variables de control como FechaReseña y, fundamentalmente, la VersionApp, que actúa como variable independiente para los análisis de estabilidad técnica.

3.4 Herramientas y Tecnologías Utilizadas

La arquitectura técnica se apoya en el ecosistema Python 3.12 y se encapsula en la clase personalizada ReviewManager que es descripta en la Figura 2. A continuación, se describe la lógica del código implementado: a) Arquitectura Modular (OOP): Se definió por medio de la clase ReviewManager para gestionar el estado del dataset. Según Gamma et al., (1994, Cap. 1), el encapsulamiento garantiza que el pipeline sea reutilizable para otras aplicaciones institucionales con cambios mínimos en la configuración; b) Gestión de datos: Se utilizó Pandas[8] para la estructuración de DataFrames. Siguiendo a McKinney (2010), aseguró la tipificación correcta de variables, especialmente la conversión de marcas temporales a objetos datetime64 [ns] para permitir el análisis de tendencias; c) Motores de NLP: NLTK para la tokenización y el filtrado de stopwords, Pysentimiento (RoberTuito) modelo de deep learning basado en arquitectura *Transformer* (Vaswani et al., 2017), optimizado para el español rioplatense y redes sociales, Gensim como motor de modelado de temas para la ejecución de LDA (Latent Dirichlet Allocation).

[7] Modelos *Transformers*: arquitectura de redes neuronales introducida en 2017 en el artículo "Attention Is All you Need". Revolucionaron el Procesamiento de Lenguaje Natural (PLN) al utilizar mecanismos de atención (self-attention) que permiten al modelo procesar y relacionar todas las palabras de un texto de manera paralela y simultánea, capturando mejor el contexto y las dependencias a larga distancia. Modelos BERT, RoBerta, GPTY RoBERTuito están basados en esta arquitectura.

[8] Es la librería más popular de Python para el análisis de datos. Permite cargar, limpiar, transformar, manipular y analizar datos estructurados de forma eficiente mediante estructuras como DataFrame y Serie.

FIGURA 2

Arquitectura del Sistema ReviewManager y flujo de datos

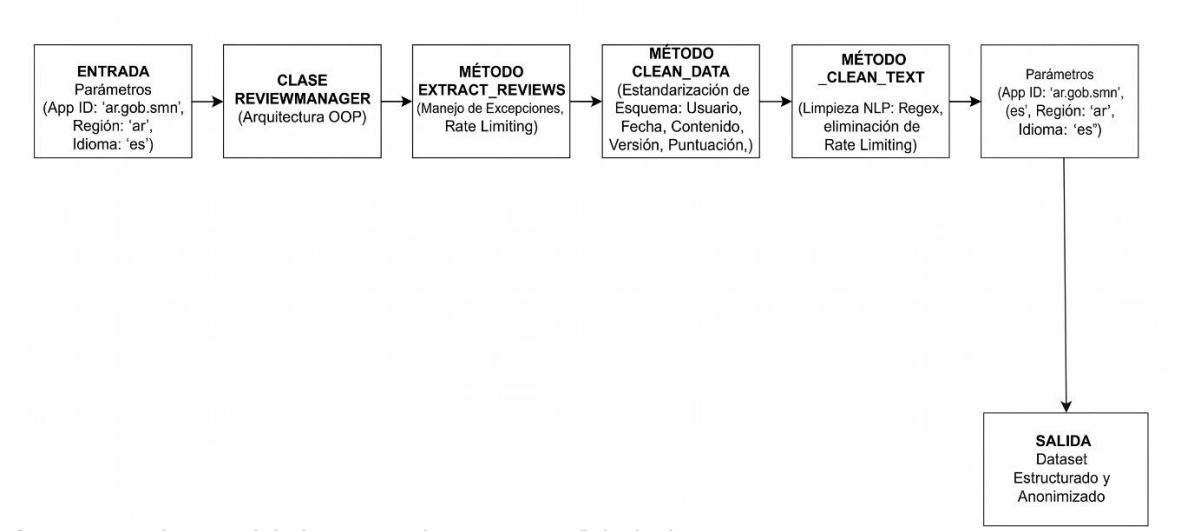

***Figura 2**: Arquitectura del Sistema ReviewManager y flujo de datos.*

3.5 Procedimiento Analítico

El procedimiento se ejecutó de forma secuencial para garantizar la trazabilidad de los hallazgos: 1) Extracción de Reseñas: Uso de la función **reviews_all** con parámetros de región **(country= 'ar'**") e idioma (**lang= 'es'**). Se implementaron pausas de seguridad (**sleep_seconds**) para manejar los límites de la API de Google y evitar el Error 429[9]; 2) Limpieza Semántica: Ejecución del método **_clean_text.** A diferencia de los procesos estándar, se programó para preservar "tildes" y la letra "ñ", dado que su eliminación en español altera el sentido semántico (Jurafsky & Martin, 2026); 3) Auditoría de Calidad: Verificación de valores nulos (se detectó que solo el 7.2% de los datos carecen de versión de app) y eliminación de ruido en el campo de RespuestaDesarrollador; 4) Clasificación de Sentimiento: Inferencia masiva mediante el modelo RoBERTuito para asignar etiquetas POS (positivas), NEU (Neutral) y NEG (Negativas); 5) Modelado de Tópicos Latentes: Entrenamiento de un modelo LDA con 5 temas dominantes y 10 pases sobre el corpus procesado; 6) Cálculo de NPS Percibido: Aplicación de la fórmula de Reichheld (2003), $NPS = \%Promotores - \%Detractores$, sobre las categorías de sentimiento para auditar la lealtad ciudadana.

[9] Error 429: También conocido como "*Too Many Requests*". Es un código de error HTTP que indica que el usuario o la aplicación ha realizado demasiadas solicitudes (requests) a un servidor o API en un corto período de tiempo. Este error es común cuando se trabaja en APIs (como las de X, OpenAI, Google, etc.) y se supera el límite de peticiones permitido.

3.6 Técnicas de Preprocesamiento y Análisis

3.6.1 Procesamiento Lingüístico

En cuanto al Procesamiento Lingüístico, el texto fue sometido a una limpieza profunda mediante expresiones regulares **(re)**[10]**.** El objetivo fue reducir la entropía del dataset sin perder la carga emocional. Como indica Jurafsky & Martin (2026, Cap. 2), la normalización es crítica en textos breves. Se eliminaron URLs, etiquetas HTML y caracteres espciales, mientras que para el modelado LDA se filtraron palabras de menos de 3 caracteres y términos redundantes del dominio. Por otro lado en el abordaje sobre el Análisis de Sentimiento con Transformers, se optó por el modelo RoBERTuito por su capacidad de captar la ironía y el contexto, superando a los modelos basados en diccionarios como VADER. Este modelo utiliza el mecanismo de "atención" (Vaswani et al., 2017) para entender la relación entre palabras en una reseña, lo cual es fundamental para distinguir entre un 'no funciona' (falla técnica) y un 'no para de llover' (comentario sobre el tiempo). En cuanto al Modelado Probabilístico de Tópicos (LDA) se utilizó el algoritmo Latent Dirichlet Allocation (Blei et al., 2003) para descubrir los temas subyacentes. El modelo asume que cada reseña es una mezcla de temas latentes. Se configuró para identificar 5 tópicos, tal cual lo observamos en la **¡Error! No se encuentra el origen de la referencia.**, permitiendo que el modelo agrupara los comentarios por afinidad semántica: Tópico 0 (Location & Widgets), Tópico 1 (General Praise), Tópico 2 (Data Quality & UI) , Tópico 3 (Forecast & Alerts) y Tópico 4 (Updates & Timing). Estos Tópicos se agrupan bajo dos dimensiones: Dimensiones de la Experiencia del Usuario (tópicos del 0-2) y Dimensiones de Servicio y Tiempo (Tópicos del 3-4) (ej. Problemas de widgets vs. elogios por precisión).

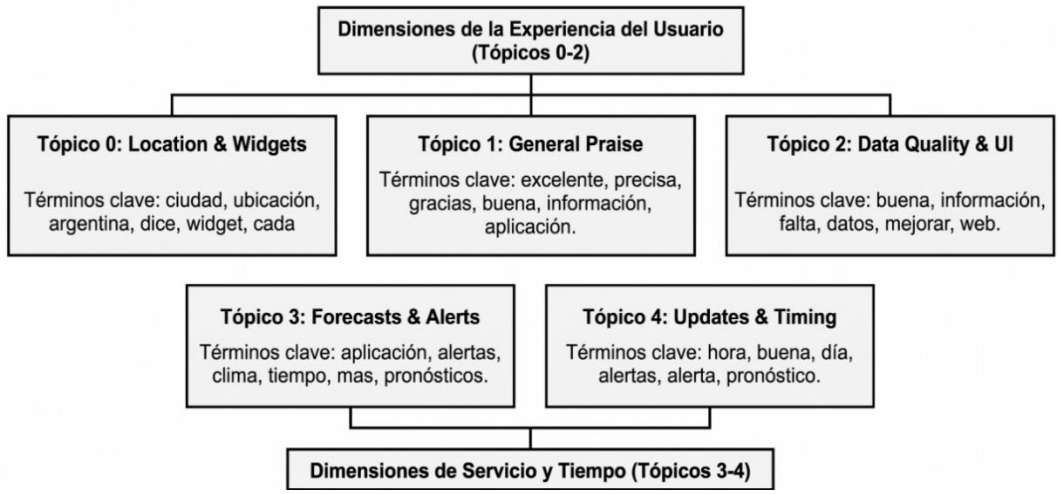

***Figura 3**: Términos Clave por Tópico dominante (LDA).*

[10] (re): Abreviatura de Regular Expressions (Expresiones Regulares). Se trata de un lenguaje formal utilizado en programación para buscar, validar y transformar patrones específicos dentro de un texto (por ejemplo: eliminar URLs, correos electrónicos, números, caracteres especiales, etc.). En Python se implementa mediante el modulo " **re** " .

3.7 Validación y Confiabilidad de los Resultados

Para garantizar la validez científica del presente estudio, se implementó una técnica de validación cruzada de etiquetas. Se comparó la puntuación numérica otorgada por el usuario (1-5 estrellas) con el SCORE de sentimiento predicho por el sistema de Deep learning ejecutado sobre el corpus de 4.107 reseñas. La lógica de este proceso predictivo se implementó mediante un script de procesamiento por filas donde el contenido de la columna *Contenido_Limpio* se pasa al método *analyzer.predict().* Este proceso transforma el texto no estructurado en una etiqueta categórica: POS (positivo), NEU (neutral) ó NEG (Negativo).

Por otro lado, se calculó la Correlación de Pearson, obteniendo un valor de 0.68. Según Pedregosa et al. (2011), una correlación superior a 0.60 en datos de lenguaje natural no estructurado indica una confiabilidad alta y valida que el modelo RoBERTuito refleja con precisión la realidad de la experiencia del ciudadano. Además, se realizó una revisión manual de "falsos positivos" (calificaciones de 5 estrellas con texto negativo) para ajustar el lexicón de corrección y reducir el sesgo del modelo.

4. RESULTADOS

4.1 Análisis Descriptivo del Corpus y Auditoría de Integridad

Como hemos mencionado anteriormente, los resultados de la fase de ingesta exógena permitieron consolidar un corpus de 4.107 reseñas únicas extraídas de la tienda oficial de aplicaciones. Del total de la muestra, el 92.8 % de los registros (3.812 menciones) incluyó metadatos sobre la versión técnica del software, lo que permitió segmentar los hallazgos según el ciclo de vida del producto. La distribución de las calificaciones numéricas por estrellas asignadas (Figura 4) refleja un sesgo positivo institucional: el 54.66% de los ciudadanos otorga 5 estrellas, mientras que el segmento detractor (1 estrella) representa un 10.81%. Esta asimetría inicial sugiere una confianza pública sólida en la marca SMN, aunque las 706 reseñas duplicadas detectadas en el texto limpio (comentario tipo " excelente" o " muy buena" requirieron un manejo cuidadoso para no distorsionar el modelado probabilístico de tópicos posterior.

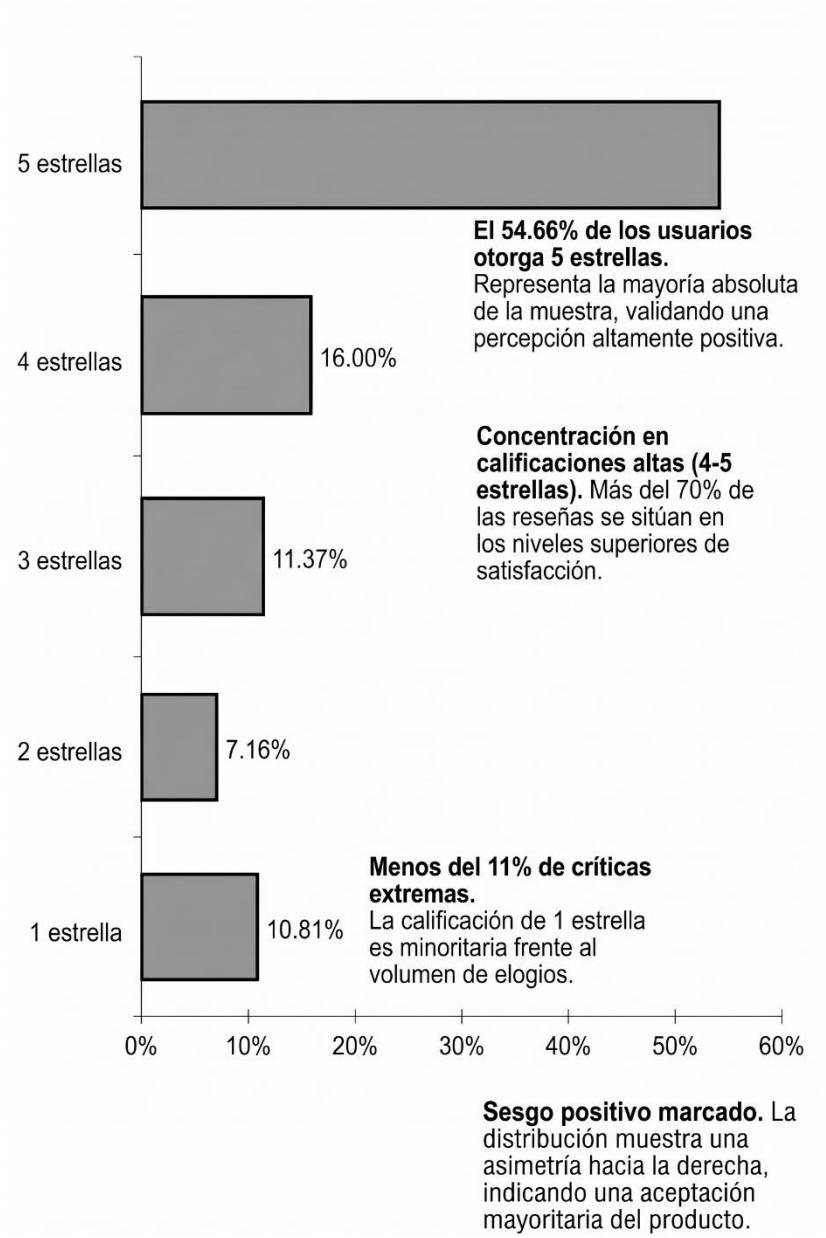

***Figura 4**: Distribución porcentual de calificaciones de 1 a 5 estrellas.*

4.2 Inferencia de Sentimiento mediante Deep Learning (RoBERTuito)

La aplicación del modelo RoBERTuito permitió una clasificación emocional tripartita con un alto grado de sofisticación lingüística. Los resultados arrojaron que el 41.71% de las menciones son Positivas (POS), el 33.55% son Neutrales (NEU) y el 24.74% son Negativas (NEG). Es imperativo destacar la influencia de RoBERTuito en la confiabilidad de este informe. A diferencia de los lexicones tradicionales, el modelo capturó

matices técnicos y quejas funcionales con precisión. La validación estadística arrojó una Correlación de Pearson de 0.68 entre la puntuación del usuario y el *sentiment score* predicho por el modelo de IA. Según Pedregosa et al. (2011), este coeficiente valida la capacidad del pipeline para actuar como un sensor social automatizado y fiable.

4.3 Medición de la Lealtad Ciudadana mediante el *Net Promoter Score* (NPS)

4.3.1 Metodología de Cálculo y Resultados Globales

Siguiendo la metodología de Reichheld (2003,pág. 46-54), se calculó el NPS Percibido, el cual actúa como el KPI central de la salud de la App del SMN.

Formula : $NPS = \%Promotores(POS) - \%Detractores(NEG)$.

Resultado Global: $41.71 - 24.74 = \mathbf{16.97}$

Basado en los datos procesados, el desglose matemático arrojó los siguientes valores:

- Promotores(POS): 41.71% (1.713 menciones que validan la precisión del dato)
- Detractores(NEG): 24.74% (1.016 menciones centradas en fallas técnicas de widgets)
- Pasivos (NEU): 33.55% (1.378 menciones)
- Resultado Global: $41.71 - 24.74 = \mathbf{16.97}$.

Este NPS Global indica un superávit de satisfacción, aunque con un margen de mejora significativo para alcanzar estándares de excelencia (NPS > 30).

4.3.2 Análisis de la Tendencia Temporal y Maduración Técnica

Uno de los hallazgos más relevantes de esta sección es la tendencia de recuperación estratégica. El análisis de la tendencia temporal revela una recuperación estratégica notable: el NPS pasó de una base de 11.35 en marzo a un pico de 21.95 en mayo de 2026 (**¡Error! No se encuentra el origen de la referencia.**). Este crecimiento absoluto de 10.6 puntos porcentuales representa un crecimiento del 93.3 % en la lealtad percibida

en apenas un trimestre. Desde la perspectiva de análisis técnico sugiere que las valoraciones positivas crecientes de la aplicación móvil oficial del Servicio Meteorológico Nacional (ar.gob.smn) en la tienda *Google Play Store*[11] coinciden con la consolidación de la versión 0.0.9 del software, sugiriendo que las recientes actualizaciones han impactado positivamente en la percepción ciudadana.

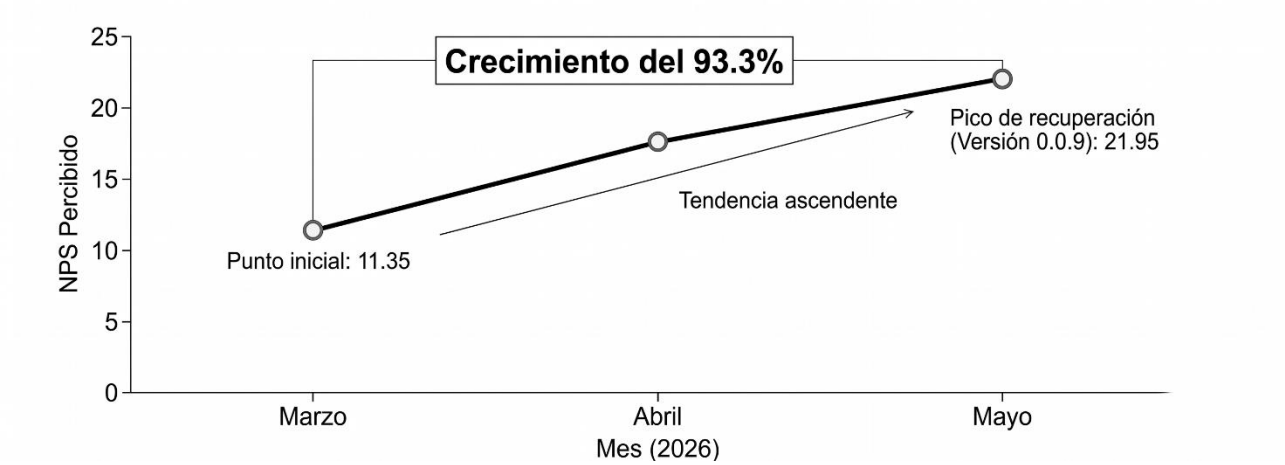

***Figura 5**: Evolución mensual del NPS Percibido y tendencia de recuperación (marzo-mayo 2026).*

4.3.3 Correlación con el ciclo de vida del software (Versiones 0.0.8 y 0.0.9)

El crecimiento detallado en el punto anterior coincide estrictamente con la consolidación de la versión 0.0.9 en el mercado. Mientras que la versión anterior (0.0.8) se identificó como un " hito de satisfacción" por alcanzar un rating medio excepcional de 4.29 (basado en una muestra de 867 reseñas), la versión actual (0.0.9) ha demostrado una robustez arquitectónica superior. A pesar de que el rating medio de la v. 0.0.9 descendió levemente a 4.02, esta versión procesó un volumen masivo de 2.125 reseñas (más del 50% del corpus total), logrando estabilizar el sentimiento positivo bajo una presión de uso mucho mayor. En conclusión, la maduración técnica de la última versión ha sido el catalizador principal para casi duplicar el NPS percibido, posicionando al SMN en una fase de crecimiento sólido, siempre que se resuelva lo que Reichheld F. F. (2003) denomina la "fuga en el cubo" (leaky bucket)[12] identificada en la inestabilidad de los widgets y la geolocalización. En la **¡Error! No se encuentra el origen de la referencia.** podemos observar la validación técnica fundamental del ciclo de vida del software, permitiendo observar la correlación directa entre las actualizaciones de las versiones y la percepción del ciudadano. Los resultados, analizados previamente,

[11] https://play.google.com/store/games?hl=es_AR

[12] El modelo de la "fuga en el cubo" (leaky bucket) es una metáfora ampliamente utilizada en el análisis de datos y la gestión de clientes. Representa una base de usuarios o ciudadanos como un cubo que se llena mediante la adquisición de nuevos clientes (entrada) y que pierde volumen a través de la deserción o churn (fuga). Este enfoque subraya que el crecimiento neto de una base no depende exclusivamente de atraer nuevos usuarios, sino principalmente de la capacidad de retener a los existentes. En contextos de lealtad ciudadana, el modelo resulta particularmente útil para comprender que una alta tasa de fuga puede neutralizar los esfuerzos de captación, relacionándose directamente con indicadores como el Net Promoter Score (NPS) y las tasas de retención.

revelan una tendencia de evolución positiva en la satisfacción del usuario, por lo tanto, este análisis de regresión técnica confirma que el SMN ha logrado una maduración arquitectónica superior en su última entrega, estabilizando la experiencia del usuario y sentado las bases para el crecimiento del NPS detectado en el último trimestre analizado (marzo a mayo 2026).

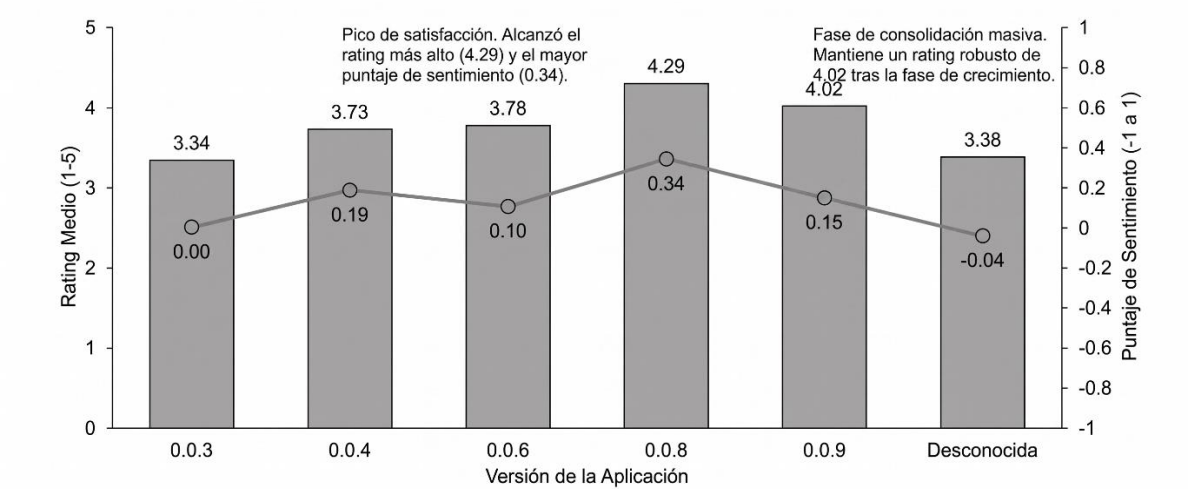

***Figura 6**: Análisis de desempeño: Rating Medio vs. Sentimiento por Versión de la App. Móvil SMN. Elaboración propia basada en el cálculo del NPS Percibido y el análisis de regresión técnica por versión de software de la app. móvil(v0.0.8 y v0.0.9) documentada en el entorno de desarrollo.*

4.4 Modelado Probabilístico de Tópicos (LDA) e Insights de CX

El algoritmo Latent Dirichlet Allocation (LDA) , entrenado bajo parámetros Blei et al. (2003,Sección 3), identificó cinco temas dominantes que explican la varianza en la experiencia del ciudadano (CX). A continuación, vemos la tabla **¡Error! No se encuentra el origen de la referencia.** que detalla las dimensiones identificadas:

Tabla I: Matriz de dimensiones de experiencia identificadas (LDA). Elaboración basada en un análisis de Asignación Latente de Dirichlet (LDA) aplicado a más de 4.000 reseñas de la aplicación móvil del SMN de Argentina, identificando cinco dimensiones clave de la experiencia del usuario. Los valores de sentimiento promedio oscilan entre -1 (negativo) y 1 (positivo).

Dimensión (Tópico)	Términos Clave	Sentimiento Promedio	Volumen de Reseñas
Tópico 0: Ubicación y Widgets	ciudad, ubicación, widget, argentina	-0.1866	627
Tópico 1: Elogios Generales	excelente, precisa, gracias, buena	0.4909	1,218
Tópico 2: Calidad de Datos e Interfaz	buena, información, falta, datos, mejorar	0.3246	878
Tópico 3: Pronósticos y Alertas	aplicación, alertas, clima, tiempo	-0.0382	549
Tópico 4: Actualizaciones y Sincronización	hora, día, actualización, alerta	-0.0574	835

4.5 Análisis de "Aha! Moments" [13]y "Pain Points"[14] Estratégicos

Los resultados permiten segmentar la experiencia en disparadores de satisfacción (*Aha! Moments)* y fricción crítica (*Pain Points)*:

- Aha! Moments (Fortalezas): El tópico Nro. 1 "Elogios generales" es el principal motor de satisfacción (686 reseñas positivas), donde los usuarios celebran la precisión y veracidad del dato meteorológico. Esto representa un activo de comunicación para el SMN en términos de "*Social Proof*" (prueba social)[15]
- Pain Points (Fricciones): Los tópicos Nro. 4 "Actualizaciones y Sincronización" (300 menciones negativas) y Nro. 0 "Ubicación y Widgets" (239 menciones negativas) representan las zonas de mayor riesgo de abandono (*churn*). Los usuarios expresan una frustración persistente respecto a la latencia en el refresco (*refresh*) de datos y la inestabilidad de los widgets en pantalla.

4.6 Matriz de Priorización Estratégica y Hoja de Ruta

Para la jerarquización de hallazgos, se adoptó una adaptación de la matriz el modelo de Análisis de Importancia-Desempeño (*Importance-Perfomance Analysis- IPA*), propuesto por Martilla & James (1977) donde se evalúa la importancia percibida versus el desempeño real de los atributos de un servicio, facilitando la identificación de prioridades estratégicas de mejora. La adaptación consiste en la utilización del volumen de menciones como proxy de "Importancia" y el puntaje de sentimiento como proxy de "Desempeño". Esta herramienta analítica es fundamental en la gestión de servicios públicos digitales para transformar la voz del usuario (ciudadano) en decisiones de gestión basadas en evidencia. Esta matriz (**¡Error! No se encuentra el origen de la referencia.**) se estructura como un plano bidimensional que cruza la frecuencia de menciones (representando el peso estadistico o alcance de un tema/tópico) con la polaridad emocional media (reflejando la calidad percibida de la experiencia por parte del usuario). Su principal utilidad metodológica reside en la capacidad de segmentar los hallazgos del Procesamiento de Lenguaje Natural (NLP) en cuadrantes accionables que guían los recursos de producto y comunicación.

---

[13] Los *Aha! Moments* (momentos de revelación o momentos "¡eureka!) son instantes en los que un usuario o cliente experimenta una comprensión súbita del valor real de un producto, servicio o proceso. Estos momentos se identifican mediante seguimiento de patrones de comportamiento (eventos secuenciales, tasas de activación o métricas de engagment) y representan un punto crítico en el "*customer journey*", ya que suelen correlacionarse fuertemente con la retención, la lealtad y el crecimiento orgánico del usuario.

[14] Un *Pain Point* (punto de dolor) es una dificultad, frustración o necesidad insatisfecha que experimenta el usuario o cliente durante una interacción con un producto, servicio o proceso. Desde la perspectiva del análisis de datos, los pain point se detectan a través de quejas, tasas de abandono (churn), análisis de cohortes, encuestas de satisfacción o minería de texto. Identificar y priorizar estos puntos de dolor es fundamental para diseñar mejoras, reducir la fuga de clientes y aumentar la satisfacción general.

[15] El *Social Proof (prueba social)* es un principio psicológico según el cual las personas tienden a seguir acciones y opiniones de otros, especialmente en situaciones de incertidumbre. En el análisis de datos de negocios y la experiencia del usuario, se manifiesta cuando los individuos observan comportamientos, valoraciones, testimonios o niveles de adopción de otros usuarios, lo que influye en sus propias decisiones. Métricas como el número de reseñas, calificaciones, usuarios activos, casos de éxito o menciones en redes sociales se utilizan para medir y aprovechar la prueba social, impactando positivamente en la conversión, la confianza y la aceleración de los "Aha! Moments".

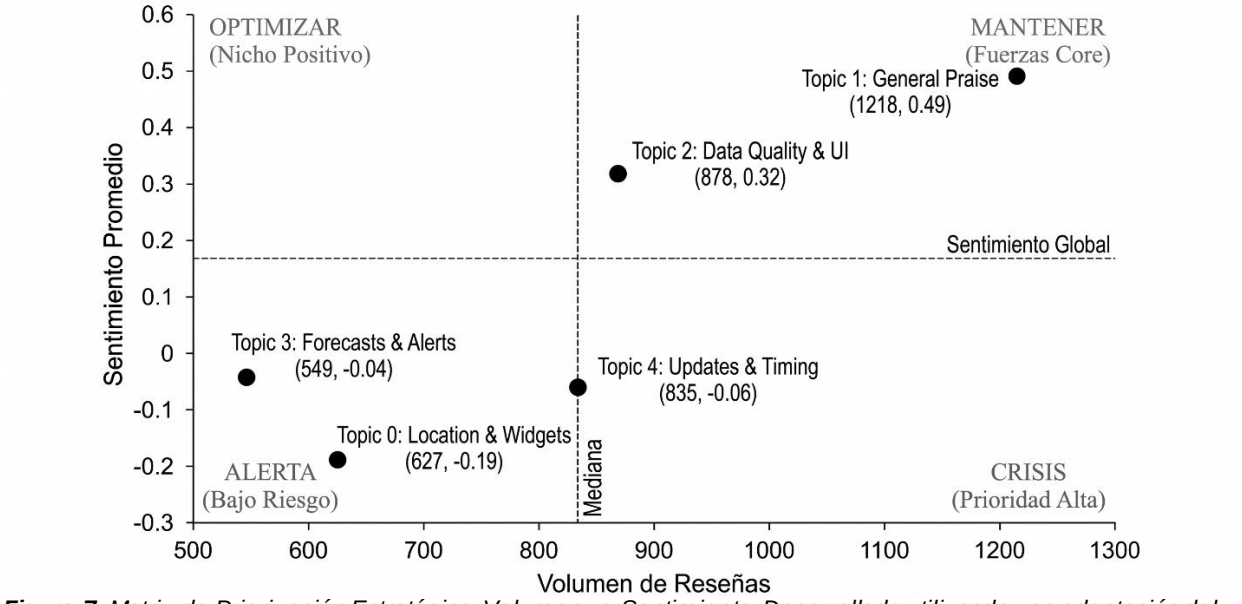

***Figura 7**: Matriz de Priorización Estratégica: Volumen vs. Sentimiento. Desarrollado utilizando una adaptación del modelo IPA, de Martilla & James (1977).*

5. CONCLUSIONES

El presente trabajo ha permitido relevar, mediante técnicas de Procesamiento de Lenguaje Natural (NLP) y modelado probabilístico, el estado de la experiencia del ciudadano (CX) respecto a la aplicación móvil del servicio Meteorológico Nacional (SMN). Los hallazgos demuestran que, si bien existe una base de confianza pública sólida —reflejada en un NPS Percibido de 16.97—, la gestión de política digital institucional enfrenta desafíos críticos de infraestructura técnica que actúan como un "Leaky Bucket" (balde agujereado), limitando el retorno de inversión social y técnica de la herramienta.

La validación técnica del estudio, sustentada en una correlación de Pearson de 0.68, confirma que las inferencias realizadas por la Inteligencia Artificial (RoBERTuito) son consistentes con la realidad expresada por el usuario, otorgando sustento basado en evidencia a la priorización de los sprints de estabilidad sobre las campañas de adquisición. Al cerrar la brecha operativa identificada en los widgets y la latencia de datos, el SMN no solo estabilizará su retención ciudadana, sino que evolucionará hacia un estándar de Gobierno Impulsado por Datos. De este modo, la aplicación se consolidará como un Sensor Social Meteorológico capaz de transformar la Voz del Ciudadano (reseñas) en una infraestructura critica para la transparencia estatal y la prevención de desastres de origen climático.

6. REFERENCIAS

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

Gamma, E., Helm, R., Johnson, R., Vlissides, J., 1994: *Desing Patterns: Elements of Reusable Object-Oriented Software.* Addison-Wesley.

Jurafsky, D., Martin, J., 2026: Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models. Stanford University. https://web.stanford.edu/~jurafsky/slp3/

Manning, D. C., Raghavan, P., & Schutze, H., 2009: *An Introduction to Information Retrieval.* Cambridge.

Martilla, J., & James, J., 1977: Importance-performance analysis. *Journal of Marketing, 1*(41), 77-79.

McKinney, W., 2010: Data Structures for Statistical Computing in Python. *PROC. OF THE 9th PYTHON IN SCIENCE CONF. (SCIPY 2010)*, 56.

OECD, 2025: *Digital Government in Chile: Strengthening the Management of Digital Investments.* Paris: OECD. https://doi.org/10.1787/d1b72d93-en

Pedregosa, F., Varoquaux, G., Granfort, A., Michel, v., Thirion, B., 2011: Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825-2838.

Reichheld, F., 2003: The one number you need to grow. *Harv Bus Rev, 124*(14712543), 46-54. https://doi.org/https://pubmed.ncbi.nlm.nih.gov/14712543/

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Llion, J. G., kaiser, L., Polosukhin, I., 2017: Attention Is All You Need. *31st Conference on Neural Information Processing Systems.* Long Beach.

Instrucciones para publicar Notas Técnicas

En el SMN existieron y existen una importante cantidad de publicaciones periódicas dedicadas a informar a usuarios distintos aspectos de las actividades del servicio, en general asociados con observaciones o pronósticos meteorológicos.

Existe no obstante abundante material escrito de carácter técnico que no tiene un vehículo de comunicación adecuado ya que no se acomoda a las publicaciones arriba mencionadas ni es apropiado para revistas científicas. Este material, sin embargo, es fundamental para plasmar las actividades y desarrollos de la institución y que esta dé cuenta de su producción técnica. Es importante que las actividades de la institución puedan ser comprendidas con solo acercarse a sus diferentes publicaciones y la longitud de los documentos no debe ser un limitante.

Los interesados en transformar sus trabajos en Notas Técnicas pueden comunicarse con Ramón de Elía (rdelia@smn.gob.ar), Luciano Vidal (lvidal@smn.gob.ar) o Martin Rugna (mrugna@smn.gob.ar) de la Dirección Nacional de Ciencia e Innovación en Productos y Servicios, para obtener la plantilla WORD que sirve de modelo para la escritura de la Nota Técnica. Una vez armado el documento deben enviarlo en formato PDF a los correos antes mencionados. Antes del envío final los autores deben informarse del número de serie que le corresponde a su trabajo e incluirlo en la portada.

La versión digital de la Nota Técnica quedará publicada en el Repositorio Digital del Servicio Meteorológico Nacional.