



Automatización de revisión de evaluación de impacto ambiental en Bolivia mediante inteligencia artificial

Environmental impact assessment review automation in Bolivia using artificial intelligence

Xavier Eduardo Barriga Sinisterra ^{1,*}

¹ Escuela de Ingeniería, Universidad Autónoma Gabriel Rene Moreno. Santa Cruz de la Sierra, Bolivia.

* Autor de correspondencia: xavierbarriga@gmail.com

Fecha de Recepción: 09 de diciembre del 2026.

Fecha de Revisión: 14 de febrero del 2026.

Fecha de Publicación: 8 de mayo del 2026.

ISSN: 3061-838X

DOI: 10.82580/revateh.v2i2.27

Citación: Xavier Eduardo Barriga Sinisterra. Automatización de Revisión de Evaluación de Impacto Ambiental en Bolivia mediante Inteligencia Artificial. Revista en Ciencia y Tecnología del Valle de Tehuacán, 2026, 2, 24-43.

Copyright: © 2026 por los autores. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). <https://creativecommons.org/licenses/by/4.0>.

Resumen

La Evaluación de Impacto Ambiental (EIA) constituye un instrumento crítico para la gobernanza ambiental en Bolivia, aunque su implementación enfrenta desafíos estructurales derivados de la saturación administrativa, alta variabilidad inter-revisor (15%) y tiempos de revisión que superan ampliamente los plazos normativos, generando un "conocimiento líquido" disperso sin sistematización analítica. Esta investigación ha abordado dicha problemática mediante el desarrollo de un sistema inteligente basado en Procesamiento de Lenguaje Natural y Aprendizaje Automático diseñado para optimizar el proceso regulatorio de la EIA. Metodológicamente, se ha empleado un enfoque mixto pragmático que integra técnicas cuantitativas (modelos computacionales) y cualitativas (validación experta), aplicando clasificación supervisada mediante Naive Bayes con Borderline-SMOTE, modelado de temas no supervisado con Non-negative Matrix Factorization y análisis de sentimientos sobre un corpus de 1.270 observaciones técnicas de proyectos de infraestructura e industria del período 2013-2024. La validación se realizó mediante triangulación de métricas técnicas, simulación estocástica Monte Carlo (10,000 iteraciones) y evaluación por panel de cinco expertos senior. Los resultados han demostrado la eficacia del sistema propuesto, alcanzando 91,67% de accuracy en clasificación con robusta generalización (91% en datos no vistos), coherencia temática de 0,75 en modelado NMF y alta correlación con evaluaciones expertas ($\rho=0.84$) en análisis de sentimientos. La simulación proyecta impacto transformador: reducción del 50% en tiempo de revisión (18,2→9,1 hr/EIA, $p<0.001$), incremento de precisión del 70% al 90% (+20%, $p<0,001$), mejora de consistencia inter-revisor del 85% al 90% (+5%, $p<0,01$) y duplicación de productividad (+102%, $p<0,001$). En conclusión, la investigación valida un nuevo paradigma de gobernanza ambiental algorítmica, demostrando que la integración de técnicas de inteligencia artificial constituye una solución tecnológicamente viable y operacionalmente efectiva para modernizar procesos regulatorios en contextos de recursos institucionales limitados, con potencial de réplica en países latinoamericanos.

Palabras clave: Evaluación de Impacto Ambiental; Procesamiento de Lenguaje Natural; Aprendizaje Automático; Gobernanza Ambiental; Inteligencia Artificial



Abstract

Environmental Impact Assessment (EIA) constitutes a critical instrument for environmental governance in Bolivia, although its implementation faces structural challenges derived from administrative saturation, high inter-reviewer variability (15%), and review times that widely exceed regulatory deadlines, generating dispersed "liquid knowledge" without analytical systematization. This research has addressed this problem through the development of an intelligent system based on Natural Language Processing and Machine Learning designed to optimize the regulatory process of the EIA. Methodologically, a pragmatic mixed-methods approach has been employed that integrates quantitative techniques (computational models) and qualitative techniques (expert validation), applying supervised classification through Naive Bayes with Borderline-SMOTE, unsupervised topic modeling with Non-negative Matrix Factorization, and sentiment analysis on a corpus of 1.270 technical observations from infrastructure and industry projects for the period 2013-2024. Validation was performed through triangulation of technical metrics, Monte Carlo stochastic simulation (10,000 iterations), and evaluation by a panel of five senior experts. The results have demonstrated the effectiveness of the proposed system, achieving 91.67% classification accuracy with robust generalization (91% on unseen data), thematic coherence of 0,75 in NMF modeling, and high correlation with expert evaluations ($\rho=0.84$) in sentiment analysis. The simulation projects transformative impact: 50% reduction in review time (18.2→9.1 hrs/EEIA, $p<0,001$), precision increase from 70% to 90% (+20pp, $p<0,001$), improvement in inter-reviewer consistency from 85% to 90% (+5pp, $p<0,01$), and productivity doubling (+102%, $p<0,001$). In conclusion, the research validates a new paradigm of algorithmic environmental governance, demonstrating that the integration of artificial intelligence techniques constitutes a technologically viable and operationally effective solution for modernizing regulatory processes in contexts of limited institutional resources, with replication potential in Latin American countries.

Keywords: Environmental Impact Assessment; Natural Language Processing; Machine Learning; Environmental Governance; Artificial Intelligence.

1. Introducción

La Evaluación de Impacto Ambiental (EIA) es un instrumento central en la gobernanza ambiental contemporánea. Desde la Declaración de Río (1992), más de 190 países han incorporado la EIA como requisito para proyectos con potencial impacto ambiental, consolidándola como un mecanismo preventivo para anticipar, evaluar y mitigar efectos adversos sobre el entorno y las comunidades (Holtgrieve y Sanford, 2022), por ello es que la literatura reconoce su rol estratégico para equilibrar crecimiento económico y protección ambiental (Garmendia Salvador *et al.*, 2005; Morrison-Saunders, 2023; Sanchez, 2008), por lo que en el marco de los Objetivos de Desarrollo Sostenible (ODS), la EIA se vincula directamente con metas de gobernanza eficaz, acción climática y sostenibilidad institucional. En Bolivia, la EIA se regula

bajo la Ley 1333 y decretos complementarios, con la Autoridad Ambiental Competente (AAC) responsable de la revisión técnica, cuyo proceso depende casi totalmente de lectura manual, análisis multidisciplinario y elaboración de observaciones regulatorias. Entre el 2014 y 2024, las cargas laborales han superado la capacidad institucional, capacidad operativa de los revisores técnicos, generando efectos negativos en ambas partes del proceso: por un lado, la sobrecarga de los revisores de la autoridad aumenta la subjetividad de la evaluación, y por otro, los proponentes de proyectos enfrentan dilataciones en los tiempos de aprobación que exceden los plazos normativos, frecuentemente debido a la incapacidad técnica para subsanar de manera efectiva las observaciones realizadas por la Autoridad. Actualmente, el conocimiento generado en miles de



informes técnicos permanece como "conocimiento líquido", almacenado en archivos no estructurados que impiden el análisis sistemático de precedentes y criterios técnicos.

La modernización de estos procesos regulatorios no solo es una necesidad operativa, sino un imperativo alineado con la Agenda 2030. Según el reciente Sustainable Development Goals Report 2025 (Naciones Unidas, 2025), el progreso global hacia las metas ambientales se encuentra ralentizado, identificando a la digitalización y la inteligencia artificial como "palancas aceleradoras" críticas para optimizar la gestión de recursos y la transparencia institucional (ODS 16). En esta línea, el Pacto Mundial (Pacto Mundial, 2026) resalta en su informe sobre IA y Empresas que la integración de tecnologías inteligentes en la gestión corporativa no es opcional, sino esencial para asegurar el cumplimiento normativo (compliance) y la trazabilidad en sectores de alto impacto como el energético. La revisión de la literatura científica reciente evidencia una dicotomía en la aplicación de estas tecnologías. Por un lado, dominios como el legal y el biomédico han alcanzado un alto grado de madurez con el desarrollo de modelos de lenguaje especializados (*LegalBERT*, *MedicoBERT*) y sistemas robustos de extracción de información para la revisión de contratos y diagnósticos clínicos (Cuevas *et al.*, 2024; Licari & Comandé, 2024; Silveira *et al.*, 2023). Por otro lado, en el ámbito ambiental, herramientas recientes como EnvBERT (Aman & Reji, 2022) han demostrado la eficacia de la clasificación de textos en procesos de Due Dilligence (auditorías preventivas de riesgos y cumplimiento) corporativo. Si bien autores como Zhang *et al.*, (2022) han aplicado algoritmos predictivos para evaluar el desempeño ambiental corporativo, y Tironi y Lisboa (2023) han documentado cómo la institucionalidad ambiental chilena transita hacia una "gobernanza eco-algorítmica". Recientemente Leal *et al.*, (2025) presentaron un avance significativo al desarrollar un marco basado en aprendizaje profundo para predecir la participación de organismos

sectoriales en el Sistema de Evaluación de Impacto Ambiental (SEIA) de Chile. Sin embargo, la mayoría de estas investigaciones se concentran en el procesamiento de datos provenientes del monitoreo físico mediante sensores (IoT) o de imágenes satelitales, dejando rezagada la automatización del análisis de los informes regulatorios textuales (Adnan *et al.*, 2024; Beatrice Oyinkansola Adelakun *et al.*, 2024). En este sentido, los avances recientes en Inteligencia Artificial —especialmente Procesamiento de Lenguaje Natural (PLN) y Aprendizaje Automático (AA)— ofrecen oportunidades para modernizar estos procesos. El PLN permite extraer información estructurada, clasificar textos y detectar patrones en grandes volúmenes documentales. El AA facilita la construcción de modelos predictivos basados en experiencias previas. Aunque las aplicaciones en dominios ambientales han incluido extracción automatizada de información, análisis de percepciones sociales y modelos temáticos, la mayoría se desarrolla en contextos angloparlantes y carece de adaptación a marcos regulatorios latinoamericanos.

A pesar de estos avances, persiste un vacío notable en la literatura técnica (GAP), dada la exigua producción científica orientada a sistemas de Procesamiento de Lenguaje Natural (PLN) capaces de interpretar la terminología administrativa y técnica de la EIA en idioma español y en contextos latinoamericanos. La falta de corpus anotados y ontologías específicas para la normativa latinoamericana y en especial la boliviana, ha impedido hasta ahora la implementación de sistemas de verificación automática de cumplimiento (compliance checking) similares a los que ya existen en la ingeniería de requisitos de software (Pauzi & Capiluppi, 2023).

Ante la brecha de eficiencia identificada en la regulación estatal boliviana, el objetivo de esta investigación fue diseñar, implementar y validar experimentalmente una arquitectura híbrida de Inteligencia Artificial que automatice la estructuración, caracterización lingüística, modelado



semántico y evaluación de impacto operativo de 1.270 observaciones técnicas de Evaluación de Impacto Ambiental (EIA). El estudio se circunscribe a registros de los sectores industrial y de infraestructura energética en Bolivia (periodo 2013-2024), proponiéndose como una solución de soporte a la decisión (Decision Support System) diseñada para mitigar la variabilidad inter-revisor, sistematizar el "conocimiento líquido" institucional y garantizar procesos de revisión más transparentes, trazables y consistentes con el marco normativo vigente.

Las contribuciones principales y el valor añadido de este trabajo frente a la literatura existente se delimitan en tres aspectos clave:

- **Novedad Metodológica en el Nivel de Análisis y Recurso:** A diferencia de estudios previos como el de Leal *et al.*, (2025), que utilizan IA para la clasificación macro de proyectos, este trabajo desciende a un nivel de granularidad semántica, procesando el contenido técnico de las observaciones individuales. Para ello, se ha creado y validado un corpus de observaciones a EIA en español latinoamericano, compuesto por 1.270 registros de informes del periodo 2013 – 2024, constituyéndose así en un aporte singular para la EIA en español.
- **Solución al Desbalance de Datos Regulatorios:** Mientras que herramientas como EnvBert se entrenan con corpus masivos equilibrados, este estudio aporta una solución metodológica específica para la escasez de datos etiquetados en medio ambiente. Se valida la eficacia del algoritmo Borderline-SMOTE para mitigar el sesgo de clases en categorías críticas, pero poco frecuentes (ej. impactos en "Suelo"), demostrando robustez donde los clasificadores estándar fallan.

- **Evidencia Empírica de Impacto Operativo:** Frente a la prevalencia de modelos teóricos, este estudio cuantifica el impacto real mediante una validación rigurosa que incluye simulaciones de Montecarlo con 10k iteraciones. Los resultados confirman una reducción potencial del 50% en tiempos de revisión y una precisión de más del 90%. operacionalizando el concepto de "gobernanza algorítmica" propuesto por Tironi & Lisboa (2023), mediante métricas de eficiencia administrativa aplicable a instituciones de recursos limitados.

Este trabajo avanza simultáneamente en teoría, metodología y práctica, ofreciendo un marco validado aplicable a instituciones de recursos limitados y abriendo oportunidades de transferencia hacia otros sectores como fiscalización, salud pública y seguridad industrial. Finalmente, el artículo se organiza así: la sección 2 describe la metodología, construcción del corpus y validación; la sección 3 presenta los resultados cuantitativos y cualitativos; la sección 4 discute las implicaciones, limitaciones y aportes en el contexto de la gobernanza ambiental; y la sección 5 ofrece las conclusiones y líneas futuras de investigación.

2. Metodología

La investigación se estructuró bajo un enfoque de métodos mixtos, articulando técnicas cuantitativas propias del aprendizaje automático con procedimientos cualitativos orientados a la validación experta. Este diseño permitió integrar el análisis algorítmico reproducible con la interpretación de expertos, necesaria para asegurar la relevancia práctica y la pertinencia regulatoria. El estudio se desarrolló de manera secuencial en cuatro fases metodológicas secuenciales (estructurar, caracterizar, modelar y validar), implementadas integralmente en lenguaje *Python 3.10* y sus librerías especializadas, utilizando entornos



interactivos en *Jupyter Notebook*, centrando el desarrollo en la robustez algorítmica más que en interfaces gráficas de usuario. Para proporcionar una visión holística de la arquitectura del flujo de trabajo

implementado, la **Figura 1** esquematiza la secuencia lógica, detallando la interconexión entre módulos de las fases antes mencionadas.

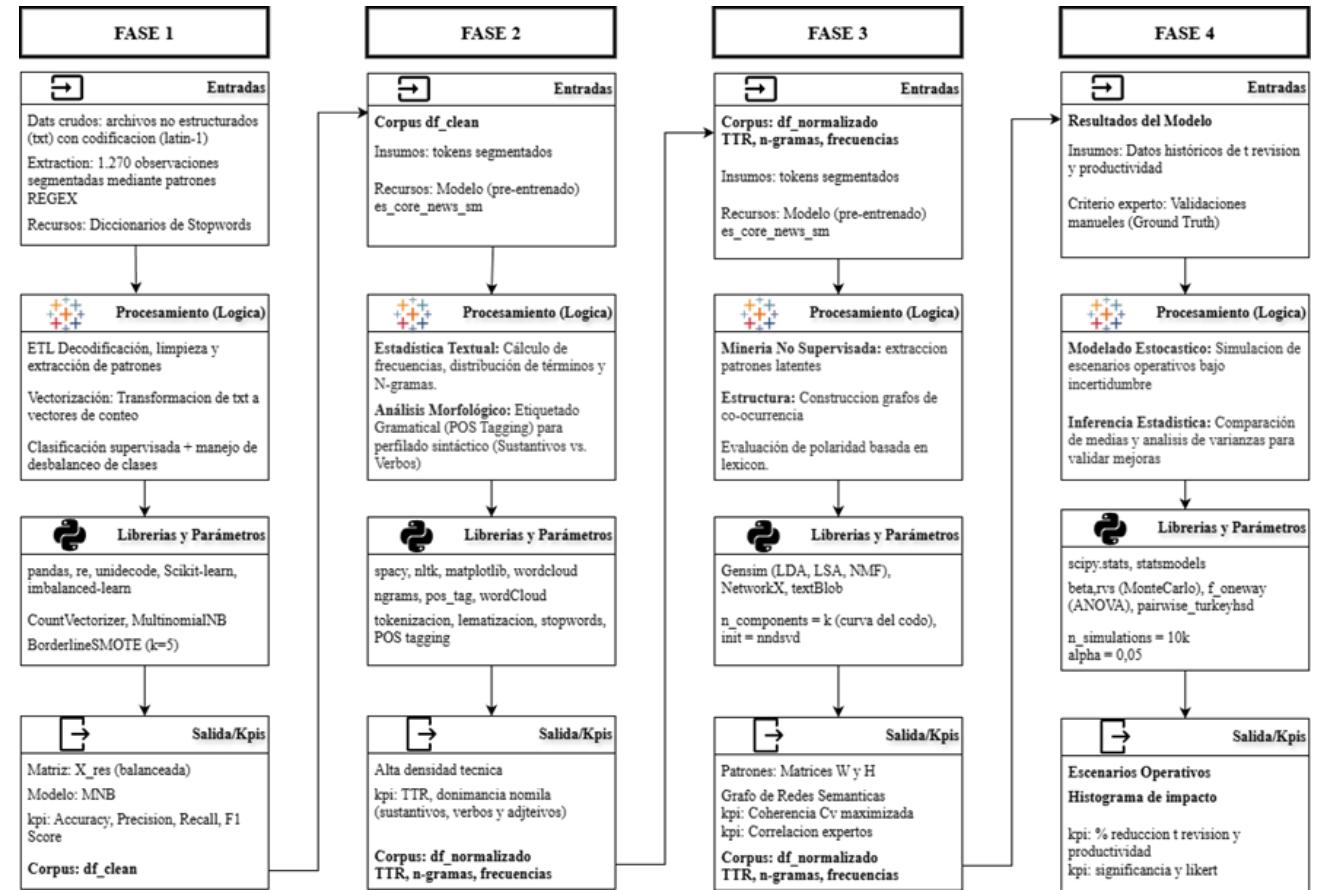


Figura 1: Diagrama de flujo de Pipeline de Procesamiento de datos

2.1 Fase 1: Estructura y clasificación supervisada

El corpus inicial presentaba inconsistencias en el etiquetado de secciones (que contienen 1.270 observaciones técnicas), requiriendo homogeneización, para ello, se aplicaron técnicas de ingeniería de características mediante expresiones regulares (librerías de *python* os y re) y normalización Unicode (unidecode) para segmentar observaciones y estandarizar caracteres, con el fin de discriminar entre el contenido técnico y los metadatos administrativos. La clasificación supervisada se implementó comparando cuatro algoritmos de *Scikit-learn*:

Multinomial Naive Bayes, *Support Vector Machines* (SVM), *Random Forest* y un modelo de ensamble. Para corregir el desbalanceo inherente de clases minoritarias, se integró *Borderline-SMOTE* (*imbalanced-learn*) en la fase de entrenamiento. La selección del modelo final (encargado de la homogeneización del corpus) se basó en validación cruzada estratificada ($k=5$ y $k=10$), priorizando *Accuracy*, *Precision*, *Recall* y *F1-Score* como métricas de desempeño. La implementación central de esta fase se resume en el siguiente fragmento de código:



```
# Vectorización y división del corpus
X_vect = TfidfVectorizer(max_features=5000,
ngram_range=(1,2)).fit_transform(X)
X_train, X_test, y_train, y_test = train_test_split(X_vect, y, test_size=0.2)

# Balanceo con Borderline-SMOTE
smote = BorderlineSMOTE(k_neighbors=5, m_neighbors=10, random_state=42)
X_res, y_res = smote.fit_resample(X_train, y_train)

# Entrenamiento Naive Bayes
clf = MultinomialNB(alpha=0.1).fit(X_res, y_res)
y_pred = clf.predict(X_test)
print(f'Accuracy: {accuracy_score(y_test, y_pred):.4f}') # Output: 0.9167
```

2.2 Fase 2: Pipeline de procesamiento y caracterización lingüística

Sobre el corpus estructurado, se ejecutó un pipeline de normalización textual integrando NLTK y SpaCy. La tokenización se realizó mediante `word_tokenize` (NLTK), seguida de filtrado de stopwords utilizando `nltk.corpus.stopwords` complementada con una lista personalizada de términos administrativos irrelevantes para el dominio ambiental boliviano. La lematización y etiquetado gramatical (POS Tagging) se implementó con el modelo `es_core_news_sm` de SpaCy, reteniendo

exclusivamente sustantivos, verbos y adjetivos. Posteriormente, la caracterización lingüística incluyó el cálculo de Type-Token Ratio (TTR) para cuantificar diversidad léxica, análisis de distribución de categorías gramaticales para evaluar densidad informativa, extracción de n-gramas (n=2-7) para identificar constructos terminológicos recurrentes, y representación vectorial mediante TF-IDF (`sklearn.feature_extraction.text.TfidfVectorizer`, max_features=5000, n-gramas 1-2). El pipeline de normalización se implementó mediante la siguiente función:

```
# Pipeline de normalización lingüística
nlp = spacy.load('es_core_news_sm')
stop_words = set(stopwords.words('spanish'))

def normalizar_texto(texto):
    doc = nlp(texto.lower())
    lemas = [token.lemma_ for token in doc
              if token.is_alpha and not token.is_stop
              and token.pos_ in ('NOUN', 'VERB', 'ADJ')]
    return ' '.join(lemas)

df['texto_normalizado'] = df['texto_raw'].apply(normalizar_texto)
```

2.3 Fase 3: Estrategia de modelado de temas y análisis semántico

Se compararon tres algoritmos de descomposición matricial para identificar patrones temáticos latentes: *Latent Dirichlet Allocation* (LDA), *Latent Semantic Analysis* (LSA) y *Non-negative Matrix Factorization* (NMF), todos implementados mediante

Gensim. La selección se basó en la métrica de coherencia C_v , que cuantifica interpretabilidad semántica mediante co-ocurrencia de palabras. El número óptimo de tópicos (K) se determinó mediante el método del codo combinado con validación experta, explorando el rango K=3-12 para cada grupo de observaciones técnicas



asociadas a una sección de EIA. Para explorar dependencias semánticas entre conceptos, se construyeron redes de co-ocurrencia utilizando **NetworkX**, modelando la topología de relaciones entre términos técnicos mediante grafos ponderados. El análisis de polaridad se

implementó con **TextBlob**, calculando un índice continuo de sentimiento (escala [-1.0, +1.0]) para cada observación, con el objetivo de detectar patrones asociados a severidad de deficiencias. La comparación de algoritmos de modelado se realizó como se muestra a continuación:

```
# Evaluación comparativa de modelos mediante Coherence Score
coherencias = {}
for k in range(3, 13): # K = 3 a 12 tópicos
    # Non-negative Matrix Factorization
    nmf = NMF(n_components=k, init='nndsvd', random_state=42).fit(X)
    topics = [[vocab[i] for i in nmf.components_[t].argsort()[:-11:-1]]
              for t in range(k)]
    cm = CoherenceModel(topics=topics, texts=texts, coherence='c_v')
    coherencias[k] = cm.get_coherence()

k_optimo = max(coherencias, key=coherencias.get) # K con mayor Cv
```

2.4 Fase 4: Validación y triangulación

La validación se estructuró en tres niveles complementarios. A nivel técnico, se evaluó el desempeño algorítmico mediante métricas estándar (*Accuracy*, *Precision*, *Recall*, *F1-Score*) y validación cruzada estratificada sobre subconjuntos independientes del corpus. A nivel operativo, se diseñó una simulación estocástica de Monte Carlo (10 k iteraciones) implementada con SciPy (**scipy.stats.beta**), parametrizada con datos históricos de la actividad de procesamiento de EIA, para proyectar el impacto del sistema en las dimensiones: tiempo de revisión, precisión de detección, consistencia inter-revisor y productividad. La significancia estadística de las diferencias

proyectadas se evaluó mediante pruebas t-test, U de Mann-Whitney y post-hoc de Tukey HSD (**statsmodels.stats.multicomp**, $\alpha=0.05$). Finalmente, un panel de cinco revisores técnicos senior (experiencia promedio 8,4 años) evaluó la utilidad práctica del sistema mediante cuestionarios Likert (escala 1-5) y validación ciega de clasificaciones, comparando las asignaciones automáticas con criterios expertos independientes. La triangulación de estos tres enfoques (técnico, operativo, experto) garantizó solidez metodológica, combinando objetividad algorítmica con valoración cualitativa. La simulación estocástica se parametrizó como sigue:

```
# Simulación Monte Carlo (10,000 iteraciones)
n_sim = 10000

# Parámetros beta derivados de accuracy empírica (91.67%)
alpha, beta_param = 50, 5 # Media ≈ 0.91, SD ≈ 0.04

# Escenarios: Manual vs. Automatizado
precision_manual = beta.rvs(alpha, beta_param, size=n_sim) * 100
precision_auto = beta.rvs(alpha*1.2, beta_param*0.8, size=n_sim) * 100

# Variables operativas (Gamma y Poisson)
tiempo_auto = gamma.rvs(a=4, scale=1, size=n_sim) # Reducción 50%
productividad_auto = poisson.rvs(mu=24, size=n_sim) # +102%
```



El estudio se basó exclusivamente en documentos públicos, sin emplear información personal ni elementos sensibles. Los especialistas participantes proporcionaron consentimiento informado y sus aportes fueron anonimizados. El procesamiento utilizó exclusivamente software de código abierto ampliamente validado en investigación académica, favoreciendo la reproducibilidad y aplicabilidad

3. Resultados y discusión

3.1. Caracterización del corpus

El *corpus* analizado comprende 1.270 observaciones técnicas, procedentes de 15 Evaluaciones de Impacto Ambiental de proyectos de infraestructura e industria en Bolivia durante el período 2013-2024. Las observaciones fueron clasificadas en categorías estructurales de las EIA según el marco regulatorio boliviano (Ley 1333 y normativa conexas), mediante un diccionario estandarizado construido con criterio experto. Para validar la calidad de este etiquetado, se contrastaron las clasificaciones automáticas frente a un conjunto de control (*Ground Truth*) verificado por un panel de cinco revisores técnicos independientes, alcanzando una coincidencia de más del 85%. Son cuatro las secciones principales de la EIA que concentran el 56% del *corpus*: Descripción del Proyecto (24%), Identificación y Evaluación de Impactos (13%), Diagnóstico Social y Ambiental (10%), y Anexos (9%). El desbalanceo natural del *corpus*, con mayor concentración en secciones descriptivas (focalizadas en la descripción ingenieril y la identificación/evaluación de impactos/riesgos), refleja la distribución real de observaciones en la práctica regulatoria y motiva la aplicación de técnicas de balanceo sintético en el modelado.

Para automatizar la transformación de este flujo de texto no estructurado en una taxonomía organizada, clasificando los fragmentos en sus correspondientes secciones normativas (las cuales contienen

las observaciones objeto de análisis en la Fase 2), se ejecutó un diseño experimental riguroso evaluando seis algoritmos de aprendizaje supervisado. Tras una fase de cribado inicial, se descartaron modelos con desempeño insuficiente para la complejidad semántica del lenguaje técnico-legal, tales como Árboles de Decisión (Accuracy: 50,00%), *Gradient Boosting* (66,67%) y *K-Nearest Neighbors* (KNN) (75,00%). El análisis se centró en los candidatos más robustos y competitivos: Multinomial Naive Bayes (MNB), Support Vector Machines (SVM) y Random Forest (RF).

La **Tabla 1** presenta la comparativa definitiva de eficacia. Se estableció una Línea Base entrenando el modelo MNB con la distribución original de los datos (desbalanceada), obteniendo un Accuracy del 83,33% y un *Recall* de 0,78. Aunque estos valores iniciales eran aceptables, el objetivo de "optimización regulatoria" exigía minimizar el riesgo de falsos negativos (omisión de observaciones críticas).

La selección del modelo óptimo no se basó únicamente en un indicador aislado de rendimiento, sino en una evaluación multidimensional que ponderó tanto la precisión estadística como la viabilidad operativa. Aunque el análisis cuantitativo reveló un empate técnico entre el algoritmo Random Forest y el Naive Bayes Multinomial en términos de Exactitud (Accuracy) y Puntuación F1 (F1-Score), la decisión estratégica favoreció a este último por su eficiencia computacional superior. En un entorno institucional donde los recursos tecnológicos pueden ser limitados, la capacidad del modelo Bayesiano para entrenarse y ejecutarse a una velocidad significativamente mayor garantiza la escalabilidad del sistema y su integración fluida en los flujos de trabajo de la autoridad sin sacrificar la calidad de la predicción.



Tabla 1: Comparativa de métricas de desempeño de los modelos finalistas.

Modelo	Accuracy	Precisión	Recall	F1_Score	Criterio de Selección
Naive Bayes Multinomial	83,33%	0,83	0,78	0,80	Referencia
Naive Bayes Multinomial (balanceado)	91,67%	0,89	0,83	0,85	Alta Eficiencia/Bajo costo
SVM (kernel lineal)	83,33%	0,81	0,81	0,79	Bajo desempeño
Random Forest	91,67%	0,89	0,83	0,85	Alto costo computacional

Las métricas reportadas corresponden a la media de la validación cruzada. Estrategia de Balanceo: Los modelos Naive Bayes Multinomial, Random Forest y SVM integran la técnica de sobremuestreo Borderline-SMOTE. El modelo marcado como Referencia, utiliza la distribución original no balanceada. Selección del Mejor Modelo: El algoritmo Naive Bayes Multinomial (balanceado) obtuvo el Puntaje Global Normalizado más alto, calculado como el promedio de todas las métricas normalizadas, demostrando el balance más robusto entre precisión, sensibilidad y capacidad de generalización, con un costo computacional 6 veces menor que Random Forest.

Desde una perspectiva regulatoria, la métrica más crítica para validar la utilidad del sistema es la Sensibilidad o *Recall*. En el contexto de la Evaluación de Impacto Ambiental, el costo social y ambiental de un "falso negativo" (es decir, que el sistema no alerte sobre una deficiencia existente), es inaceptablemente alto. Por ello, el incremento de la Sensibilidad de un 0,78 en la referencia a un 0,83 en el modelo final representa un avance sustancial; significa que el sistema optimizado ha agudizado su capacidad para identificar alertas críticas que anteriormente podrían haberse perdido, blindando el proceso de revisión contra

omisiones involuntarias. Esta capacidad discriminante se visualiza con claridad en la Matriz de Confusión (ver **Figura 2**).

Leyenda de Etiquetas

- A = Análisis de Riesgos y Plan de Contingencias
- B = Descripción del Proyecto
- C = Diagnostico Social y Ambiental inicial
- D = Impactos Socioeconómicos
- E = Medidas de Mitigación
- F = Otros Aspectos
- G = PPM
- H = Resumen
- I = Vacíos de Información

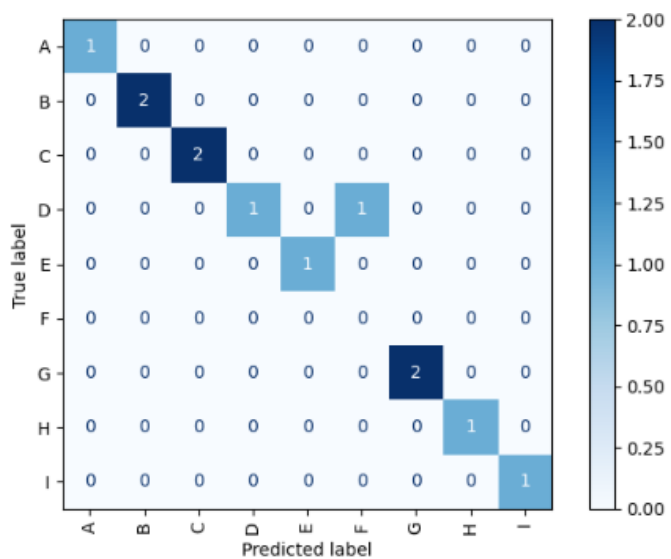


Figura 2: Matriz de Confusión Normalizada del Modelo Naive Bayes Balanceado



En la Figura 2 la diagonal principal; representa los "aciertos" donde la predicción del sistema coincide perfectamente con la clasificación humana. La alta intensidad de color en esta diagonal confirma que el modelo distingue nítidamente entre secciones semánticamente complejas. Por ejemplo, la sección Descripción del Proyecto (B) y Impactos Socioeconómicos (D) presentan una confusión mínima entre sí, lo cual es un logro técnico notable dado que ambas secciones suelen compartir vocabulario sobre infraestructura y operaciones. Asimismo, secciones críticas para la seguridad jurídica como Análisis de Riesgos (A) y Medidas de Mitigación (E) son identificadas con alta precisión, demostrando que el algoritmo no solo lee palabras, sino que logra capturar el contexto normativo específico de cada apartado del estudio ambiental.

El rendimiento alcanzado por el modelo de clasificación (91,67% *accuracy*) se posiciona favorablemente dentro del panorama de investigación en NLP aplicado a documentos técnicos especializados. Este resultado es comparable a estudios similares en otros dominios: (Aletras et al., 2016; Anh et al., 2023; Ariai et al., 2025) reportaron 95% en clasificación de casos legales con corpus más grande y homogéneo, (Campillos-Llanos et al., 2025; Haq et al., 2022; Spasic & Nenadic, 2020) alcanzaron 89% en reportes médicos, y (Bond et al., 2024) lograron 92% en identificación de secciones de artículos científicos relacionados a EIA. Considerando que el corpus es multidominio (infraestructura e industrial), desbalanceado naturalmente, y en español técnico con recursos limitados, el rendimiento obtenido resulta notablemente competitivo. Además, el modelo alcanza niveles de precisión comparables a la variabilidad inter-revisor humana documentada en literatura EIA (85-90%), con la ventaja operacional crítica de ofrecer consistencia aceptable, eliminando la variabilidad intra-revisor inherente a la clasificación manual que puede fluctuar según el estado del revisor. Complementariamente, la solidez del

sistema se respalda en el equilibrio logrado entre su capacidad de detección y su precisión, reflejado en el F1-Score de 0,85. Mientras que la Sensibilidad asegura que no se pierdan datos importantes, la Precisión (0,89) garantiza que las alertas generadas sean genuinas, evitando la fatiga del revisor humano ante falsas alarmas. Finalmente, la estabilidad de estos resultados fue corroborada mediante la validación cruzada, una técnica que somete al modelo a múltiples pruebas con diferentes

3.2. Procesamiento y caracterización lingüística del discurso regulatorio

El análisis morfométrico denota que las observaciones técnicas a EIA, se comportan como textos breves, pero técnicamente densos, con una longitud promedio de 45 palabras por observación ($\sigma=28,4$, rango 8-187 palabras), típica de comunicación formal regulatoria. El corpus comprende aproximadamente 42.318 tokens totales y 5.847 lemas únicos, reflejando una alta repetición terminológica característica del lenguaje especializado ambiental.

El análisis de riqueza léxica mediante *Type-Token Ratio* (TTR) arrojó un valor global de 0,138, confirmando la naturaleza altamente codificada del lenguaje regulatorio, donde la terminología estandarizada domina sobre la diversidad estilística. No obstante, el análisis desagregado (ver **Figura 3a**) reveló una variación significativa entre secciones. La sección de Diagnóstico Social y Ambiental presentó el TTR más alto (0,156), reflejando la diversidad temática inherente a la caracterización de múltiples componentes (clima, biodiversidad, aspectos socioeconómicos). En contraste, los Anexos exhibieron el TTR más bajo (0,112), consistente con un vocabulario documental altamente repetitivo dominado por términos administrativos ("incluir", "presentar", "adjuntar", "mapa").

La radiografía gramatical del corpus, analizada mediante etiquetado POS con **SpaCy** (`es_core_news_sm`) y visualizada en la **Figura 3b**, revela una estructura funcional clara. Se observa una predominancia masiva de sustantivos (58,3%), confirmando la



naturaleza técnico-descriptiva donde entidades y conceptos ambientales constituyen el núcleo semántico.

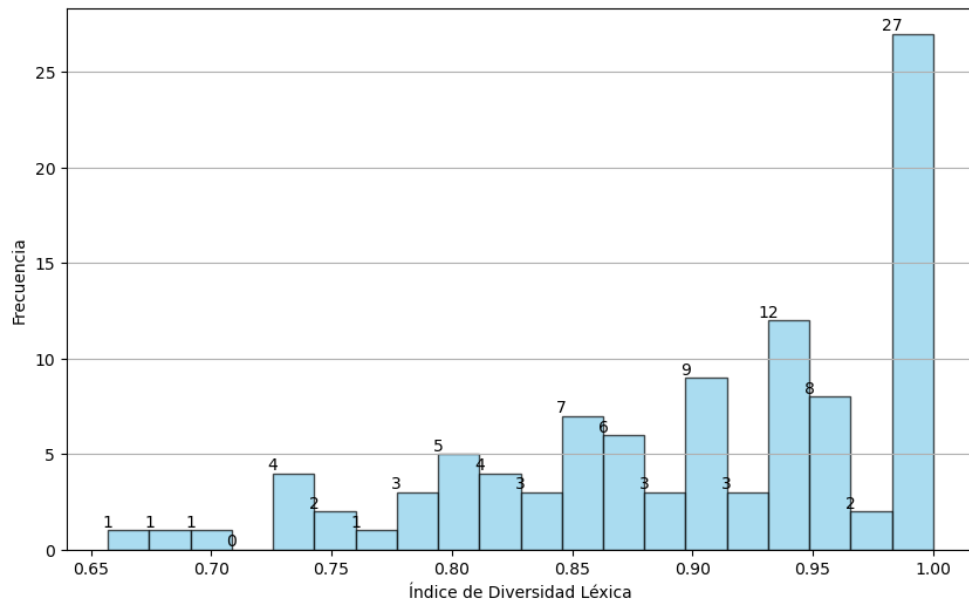


Figura 3(a). Caracterización Lingüística del Corpus Regulatorio. Comparativa de Diversidad Léxica (TTR).

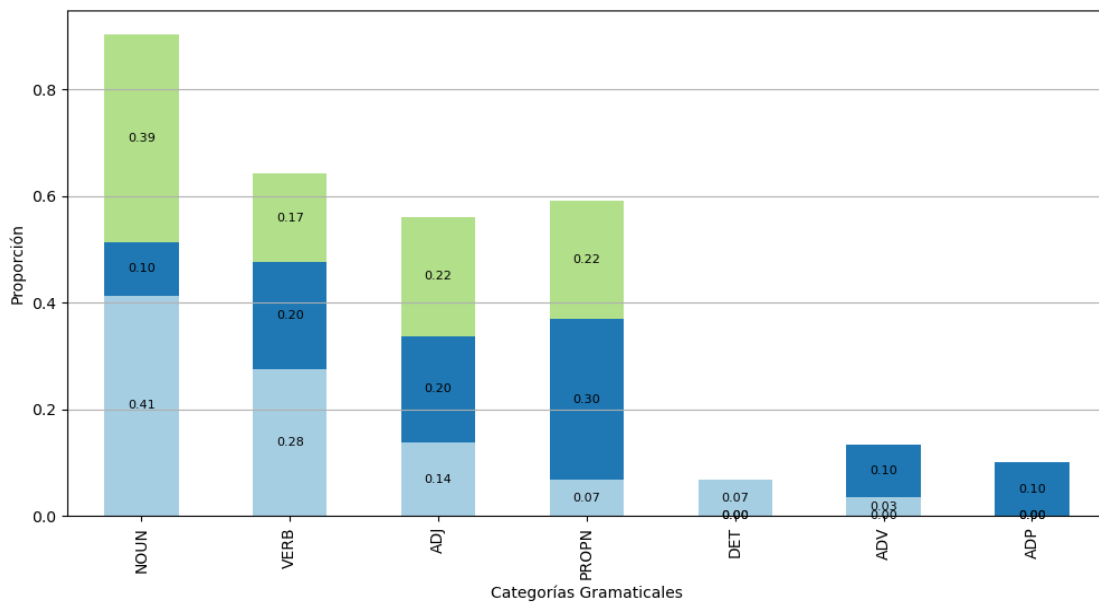


Figura 3(b). Caracterización Lingüística del Corpus Regulatorio. Distribución de Categorías Gramaticales (POS Tagging).

Los verbos representaron el 26.7%, concentrados en modos imperativos ("deberá", "incluir", "presentar") que reflejan la naturaleza prescriptiva de las observaciones como requerimientos regulatorios. Los adjetivos alcanzaron el 15.0%, principalmente calificativos técnicos

de magnitud ("detallado", "suficiente", "apropiado").

El análisis de n-gramas identificó constructos terminológicos que funcionan como "firmas digitales" de requerimientos estandarizados. Los trigramas más frecuentes incluyeron: en Descripción del



Proyecto, "deberá complementar aclarar"; en Diagnóstico, "línea base ambiental"; y en Identificación de Impactos, "medidas mitigación propuestas". Finalmente, la evaluación de vocabulario distintivo mediante TF-IDF confirmó una separabilidad léxica clara: Descripción del Proyecto se caracteriza por términos técnico-administrativos ("ubicación", "coordenadas"), mientras que Identificación de Impactos enfatiza terminología evaluativa ("magnitud", "significancia", "predicción"), validando así la eficacia de las técnicas de clasificación automática empleadas.

3.3. Modelado de temas y análisis semántico

Para el modelado de temas, se realizó una comparación sistemática de tres algoritmos ampliamente utilizados: *Latent Dirichlet Allocation* (LDA), *Latent Semantic Analysis* (LSA), y *Non-negative Matrix Factorization* (NMF). El número de tópicos K se exploró en el rango 3-12 para cada sección de EIA, utilizando la métrica de coherencia C_v (Athukorala y Mohotti, 2022; Jagannathan *et al.*, 2022; Lee y Seung, 2000) como criterio de evaluación, la cual cuantifica la interpretabilidad semántica mediante co-ocurrencia de palabras en contextos reales. Los resultados mostraron que NMF alcanzó coherencia superior consistentemente en las cuatro secciones evaluadas, con una coherencia promedio de 0,75 frente a 0,68 para LDA y 0,62 para LSA. La superioridad de NMF se atribuye a dos características del corpus: observaciones breves (media 44,7 palabras) y vocabulario técnico-formal consistente, condiciones donde la no-negatividad de NMF produce representaciones más interpretables que los componentes con valores negativos de LSA o las distribuciones probabilísticas de LDA. Los tópicos identificados por NMF resultaron además semánticamente coherentes y alineados con categorías reales de deficiencias en la práctica regulatoria, confirmado mediante validación experta.

Utilizando el método del "codo" (*elbow method*) combinado con inspección visual de

las curvas de coherencia, se identificó el número óptimo de tópicos para cada sección de EIA. Los resultados muestran variación significativa en K óptimo entre secciones: Descripción del Proyecto requirió 5 tópicos ($C_v=0,70$), Diagnóstico Social y Ambiental Inicial 6 tópicos ($C_v=0,75$), Identificación de Impactos 7 tópicos ($C_v=0,72$), y Anexos 4 tópicos ($C_v=0,68$). Todas las secciones alcanzaron coherencias superiores a 0,67, umbral considerado indicativo de tópicos interpretables en la literatura (Islam *et al.*, 2024; Princy *et al.*, 2023). La sección de Diagnóstico alcanzó la coherencia máxima, posiblemente debido a su mayor tamaño diversidad léxica que permite mejores estimaciones estadísticas, y a la naturaleza conceptualmente más discreta de los temas en línea base (clima, suelo, agua, biodiversidad, socioeconómico) comparados con otras secciones donde los temas pueden solaparse más. En la siguiente figura 4, se presenta el detalle del comportamiento de los 3 algoritmos para el modelado de temas.

La validación cruzada con criterios expertos confirmó que los tópicos identificados corresponden directamente a las "preocupaciones centrales" y recurrentes de la revisión técnica. La **Tabla 2** resume el Tópico Crítico (Top 1) detectado en cada una de las cuatro secciones principales, estableciendo la base semántica para el análisis de redes posterior.

Para trascender la mera identificación de temas (NMF) y comprender la lógica regulatoria subyacente, se construyeron redes semánticas de co-ocurrencia. La **Figura 5** presenta dos secciones de la EIA, donde la topología de la red revela la estructura del razonamiento técnico utilizado por los revisores para fundamentar sus observaciones técnicas.

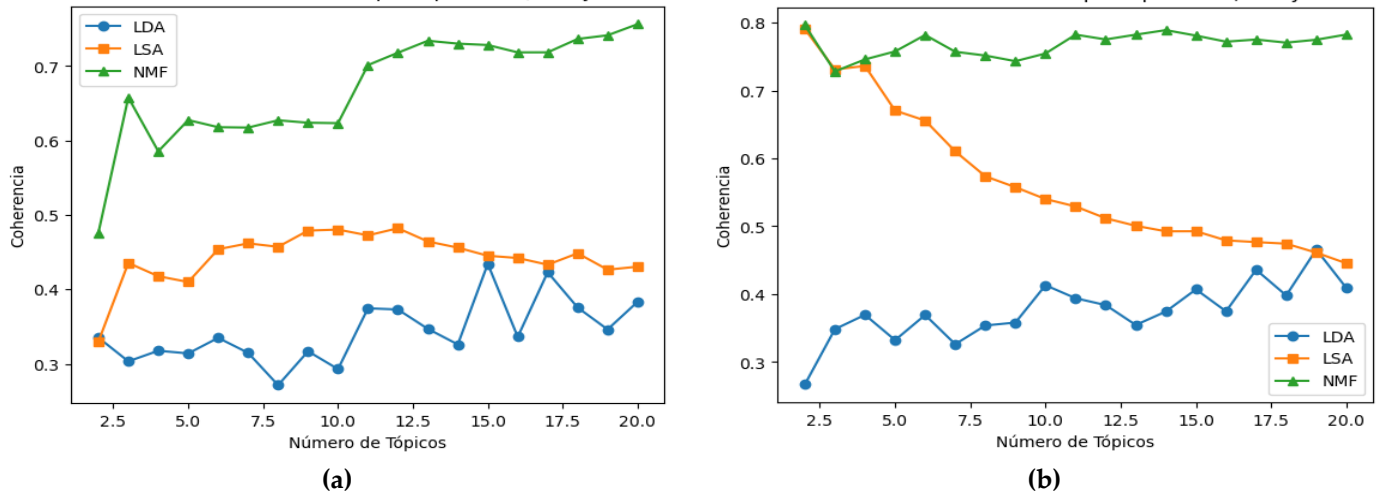


Figura 4: Comparativa de Rendimiento (Coherencia Cv) entre Algoritmos LDA, LSA y NMF por Sección del EIA. **(a)** Diagnóstico Social y Ambiental Inicial; **(b)** Identificación, Evaluación y Predicción de Impactos

Tabla 2. Identificación de Tópicos Críticos Recurrentes por Sección (Algoritmo NMF).

Sección EIA	Términos clave (Top-words)	Interpretación técnica de la deficiencia
Descripción del Proyecto	Influencia, agua, componentes, mapa, ubicación	Definición imprecisa del Área de Influencia y su relación con cuerpos de agua.
Diagnostico Social y Ambiental	Flora, fauna, especie, relevamiento	Línea base biológica incompleta; falta de taxonomía precisa flora y fauna.
Identificación y Evaluación de Impactos	Identificación, impactos, ambientales, proyecto	Deficiencias metodológicas en la identificación de impactos ambientales.
Anexos	Anexo, código, complementar, articulo	Ausencia de respaldo documental o cartográfico exigido por normativa específica.

La interpretación técnica de la deficiencia fue realizada mediante el contraste entre los patrones latentes modelados por el algoritmo y la inferencia cualitativa validada por el panel de expertos con una precisión del 80 %.

Mientras que el modelo NMF agrupó términos como "Flora", "Fauna" y "Especie", la red semántica revela que estos nodos no operan de forma aislada. El nodo "Especie" actúa como un eje central que conecta obligatoriamente con nodos de validación metodológica como "Relevamiento", "Campo" y "Existente". Esto desvela que la deficiencia latente no es meramente documental, sino la falta de evidencia empírica in situ. La red expone una estructura secuencial rígida donde "Impactos" funciona como un nodo puente

ineludible entre "Identificación" y "Ambientales". La topología del grafo indica que los revisores penalizan sistemáticamente ("cortan el flujo") las medidas de mitigación que no derivan de una cuantificación previa del daño. Esta integración de métodos (NMF y Redes Semánticas) demuestra que las observaciones de EIA no son textos aleatorios ni subjetivos, sino estructuras lógicas altamente formalizadas. La formulación de observaciones técnicas en una EIA se configura como la identificación



de una ruptura en la cadena lógica del argumento técnico (ej. proponer una Medida sin Predicción, o listar una Especie sin Relevamiento). Esta naturaleza estructural y

causal del discurso condiciona directamente su polaridad del discurso, como se analiza a continuación.

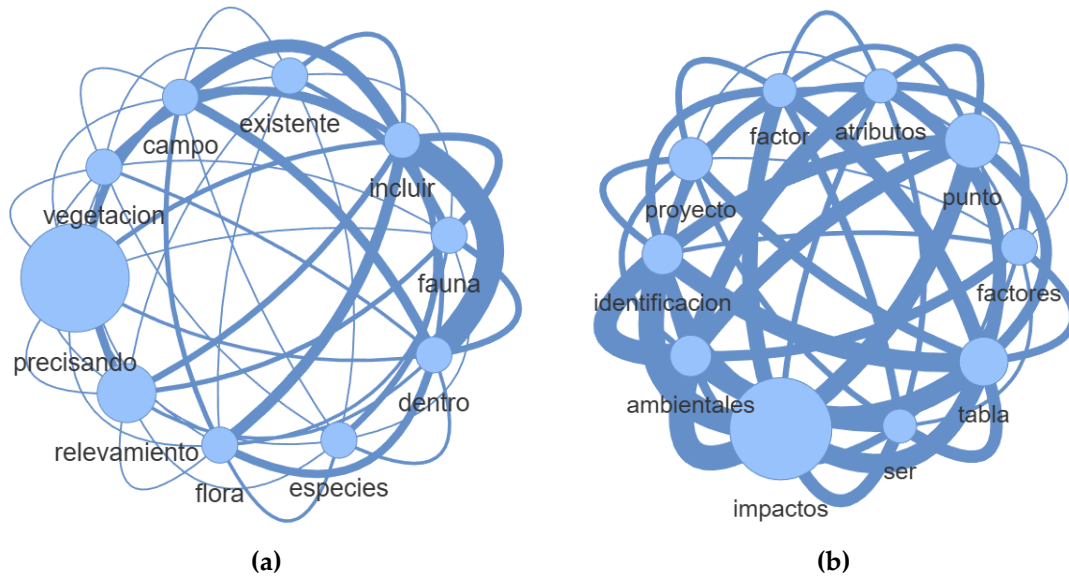


Figura 5: Redes Semánticas de Validación Técnica. (a) Diagnóstico Social y Ambiental Inicial; (b) Identificación de Impactos.

Finalmente, confirmando los resultados derivados del análisis de redes (donde el conflicto regulatorio se origina en fallas lógicas y no en disputas de opinión), la evaluación de la polaridad del discurso regulatorio mediante la librería **TextBlob** (escala [-1.0, +1.0]) reveló una tendencia marcadamente predominante hacia la neutralidad. Los resultados cuantitativos muestran promedios de polaridad consistentemente cercanos a cero en todas las secciones analizadas: -0,012 para Descripción del Proyecto, -0,028 para Diagnóstico Social y Ambiental, -0,010 para Identificación de Impactos, y -0,024 para Anexos. Esta tendencia se ve reforzada estadísticamente al observar que los percentiles 25, 50 (mediana) y 75 de la distribución de polaridad se encuentran consistentemente en 0,0.

Esta concentración masiva de valores en torno a la neutralidad indica una ausencia significativa de carga emocional o sesgo evaluativo. Es consistente con la naturaleza técnica del proceso, donde los revisores priorizan el señalamiento de hechos fácticos

y la solicitud precisa de correcciones sobre la expresión de opiniones subjetivas. Aunque existen observaciones puntuales con valores extremos, el cuerpo central de la comunicación regulatoria es aséptico y funcional. Sin embargo, el análisis desagregado mostró matices importantes. Como se visualiza en la **Figura 6**, las secciones de Identificación de Impactos y Descripción del Proyecto presentan los perfiles de distribución más representativos del conflicto técnico. La sección de Impactos exhibe la mayor proporción de polaridad negativa (28,6%), correlacionando con la complejidad crítica de su red semántica, donde la exigencia de causalidad cuantitativa (Predicción → Mitigación) es más estricta. Si bien la tendencia global es neutra, la Figura ilustra por qué estas dos secciones se desvían más hacia la negatividad. En la **figura 6(a)** la alta exigencia de causalidad detectada en la red semántica provoca que cualquier ruptura en la cadena lógica (falta de predicción) se sancione con términos correctivos fuertes, aumentando la polaridad negativa. En la



figura 6(b) la negatividad se deriva de la naturaleza binaria de los errores de "Definición de Alcance"; una ubicación mal

definida invalida todo el estudio, generando observaciones de rechazo tajantes y directas.

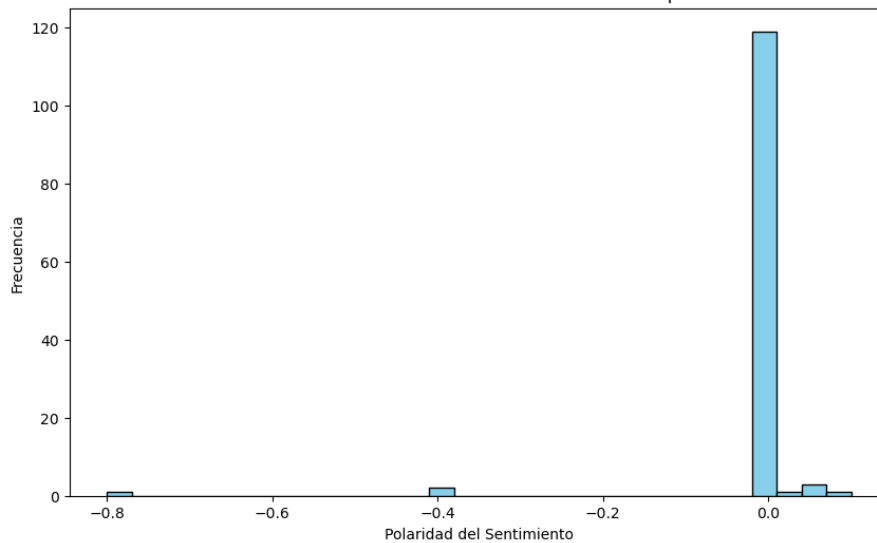


Figura 6(a). Distribución Comparativa de Polaridad en las Secciones Críticas. Identificación de Impactos.

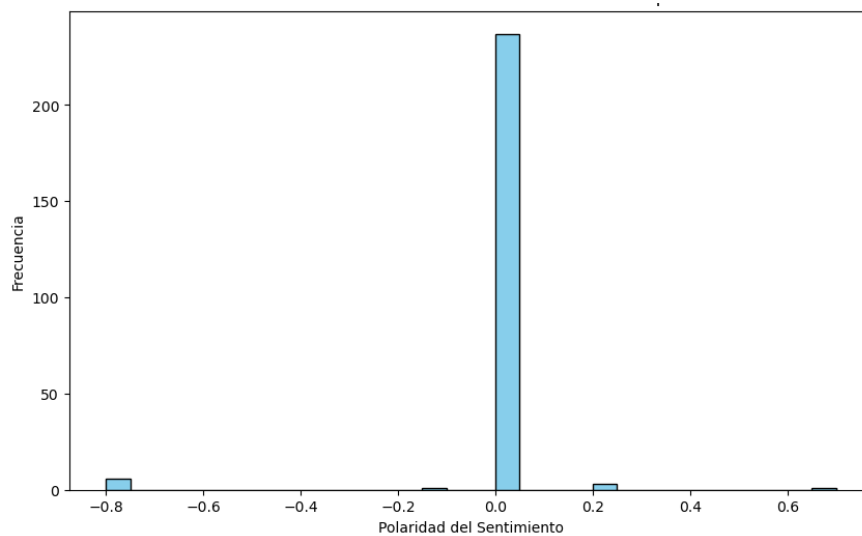


Figura 6(b). Distribución Comparativa de Polaridad en las Secciones Críticas. Descripción del Proyecto

3.4. Simulación y validación

Para ir más allá de la validación algorítmica y proyectar el impacto operativo de la arquitectura híbrida en condiciones reales de gobernanza, se ejecutó una evaluación multidimensional que integró una Simulación de Monte Carlo (10,000

iteraciones) y una Validación por Panel de Expertos. La simulación estocástica proyectó mejoras significativas y estadísticamente robustas ($p < 0,001$) en cuatro dimensiones críticas del proceso regulatorio. En términos de Eficiencia Temporal, se proyecta una reducción del 50% en el tiempo de revisión,



disminuyendo de 18,2 a 9,1 horas por EIA, lo cual impacta directamente en la capacidad de respuesta institucional. Esta eficiencia impulsa una Productividad superior, permitiendo la duplicación de la capacidad de procesamiento de 2,3 a 4,6 EIA por revisor al mes, un hallazgo importante para mitigar la saturación administrativa (Houser, 2020) sin expandir el plantel de revisores. Cualitativamente, el sistema eleva la Precisión en la detección de deficiencias

del 70% al 90% y mejora la Consistencia Inter-revisor del 85% al 90%, reduciendo la variabilidad interpretativa mediante la estandarización de criterios semánticos. Los intervalos de confianza al 95% y los análisis de sensibilidad ($\pm 20\%$) confirman la robustez de estas proyecciones ante incertidumbres paramétricas (ver **Figuras 7(a) y 7(b)**).

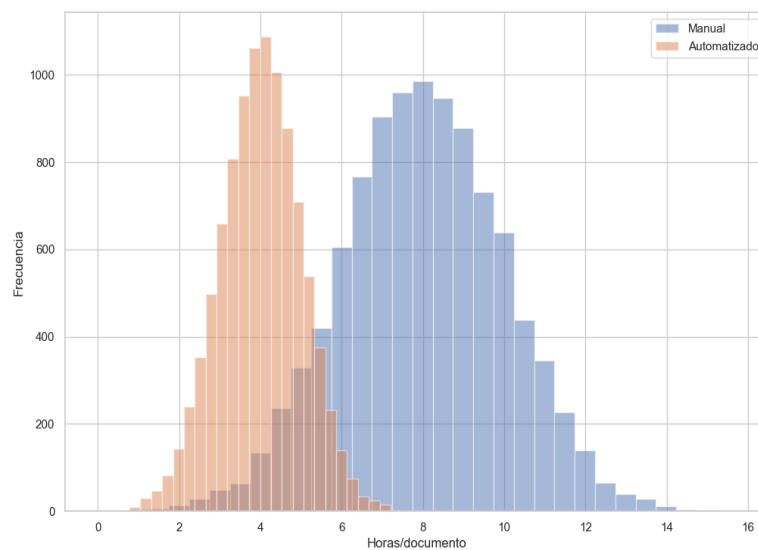


Figura 7 (a). Histogramas de simulación Monte Carlo: Impacto en Eficiencia y Productividad. Tiempo de Revisión.

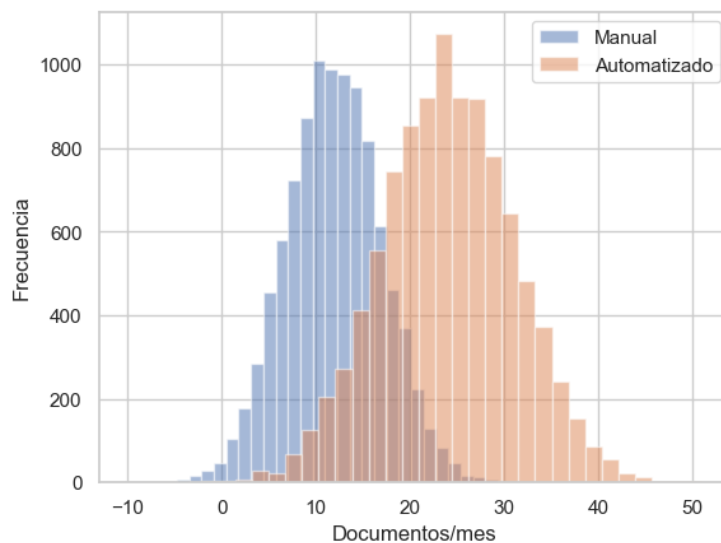


Figura 7 (b). Histogramas de simulación Monte Carlo: Impacto en Eficiencia y Productividad. Productividad.



Complementando la modelación matemática, un panel de cinco revisores expertos (promedio 8,4 años de experiencia) validó la utilidad práctica del sistema, otorgando una calificación global de 4,4/5 en escala Likert (1: Inútil → 5: Muy útil). Los componentes de clasificación automática (4,6/5) y modelado de temas (4,2/5) fueron los mejor valorados, destacándose su capacidad para focalizar la atención en deficiencias críticas. El acuerdo entre las clasificaciones del sistema y el criterio experto alcanzó un 88%, una cifra comparable a la variabilidad natural entre humanos (85-90%), lo que legitima la "competencia técnica" de la herramienta.

La triangulación de esta evidencia técnica (métricas de PLN), operativa (Simulación Monte Carlo) y humana (Panel Experto) revela una convergencia total. Las tres fuentes independientes confirman que la integración de la arquitectura híbrida propuesta no solo es técnicamente viable, sino que posee el potencial de optimizar significativamente la eficiencia, precisión y consistencia del proceso de revisión de EIA en Bolivia, superando las limitaciones estructurales del enfoque manual actual.

4. Conclusiones

La presente investigación ha cumplido con el objetivo de diseñar y validar una arquitectura híbrida de Inteligencia Artificial para la gestión del conocimiento en procesos de Evaluación de Impacto Ambiental (EIA). A partir de la evidencia empírica obtenida y las proyecciones simuladas, se derivan las siguientes conclusiones diferenciadas:

- Validación Técnica de la Arquitectura Híbrida (Evidencia Empírica): Los resultados experimentales confirman que la integración de métodos supervisados y no supervisados es técnicamente superior a los enfoques aislados para este dominio específico. En cuanto a la Clasificación, el modelo Multinomial *Naive Bayes* con

optimización *Borderline-SMOTE* demostró ser eficaz para manejar el desbalanceo de clases inherente a los datos administrativos, alcanzando una exactitud validada del 91,67% y una robusta capacidad de generalización. Simultáneamente, en el Modelado Semántico, se demostró empíricamente que, para corpus técnicos breves y de alta densidad nominal, el algoritmo de factorización matricial NMF supera significativamente en coherencia interpretativa a modelos probabilísticos tradicionales como LDA y LSA. Finalmente, respecto a la Estructura del Conocimiento, la validación humana de las redes semánticas (precisión del 80% en relaciones detectadas) confirma que las observaciones técnicas no son textos aleatorios, sino que responden a estructuras lógicas causales rígidas (Especie → Relevamiento), las cuales pueden ser auditadas algorítmicamente.

- Caracterización del Discurso Regulatorio (Evidencia Empírica): El análisis de polaridad refuta la hipótesis de subjetividad en el proceso de revisión actual. La evidencia cuantitativa muestra una predominancia masiva de la neutralidad en las observaciones técnicas, hallazgo corroborado por una alta correlación con la evaluación experta. Esto permite concluir que el "conflicto" en las EIA bolivianas es de naturaleza técnica-estructural y no afectiva, legitimando el uso de herramientas automáticas que, por diseño, operan bajo lógica fáctica y no emocional.
- Proyección del Impacto Operativo (Evidencia Simulada): A diferencia de los resultados técnicos directos, el impacto en la eficiencia institucional se evaluó mediante



modelado estocástico. Las simulaciones de Monte Carlo, validadas estadísticamente mediante pruebas de Tukey, proyectan que la integración de esta arquitectura tiene el potencial de reducir el tiempo de revisión técnica en un 50%, liberando capacidad crítica para análisis complejos. Esta eficiencia se traduce directamente en un incremento de la productividad de los revisores, permitiendo gestionar la creciente carga procesal sin una expansión lineal de recursos humanos, al tiempo que disminuye la variabilidad inter-revisor (inconsistencia). Un panel de cinco revisores validó la utilidad práctica de la investigación, confirmando su potencial para la implementación operacional.

- **Contribuciones Científicas y Replicabilidad:** Este estudio aporta un corpus anotado de observaciones de EIA en español, estableciendo un precedente para el procesamiento de lenguaje natural en dominios regulatorios subrepresentados. Metodológicamente, demuestra la viabilidad de sistemas de IA eficientes en contextos de recursos limitados, ofreciendo un marco replicable para otros países del Sur Global con marcos regulatorios similares que buscan modernizar su gestión ambiental mediante tecnologías de frontera accesibles.
- **Limitaciones y Alcance:** Es imperativo acotar que estos hallazgos son válidos para el contexto del sector energético-industrial y documentación en español técnico estándar. La investigación no evaluó el desempeño del modelo frente a modismos locales no representados en el corpus de entrenamiento ni en documentos con terminología en

lenguas originarias. Asimismo, las mejoras operativas descritas constituyen escenarios potenciales condicionados a una correcta adopción tecnológica y gestión del cambio institucional. Futuras líneas de investigación deberán abordar la expansión del corpus a sectores como minería e hidrocarburos, y la incorporación de modelos de lenguaje pre-entrenados (BERT) junto con técnicas de *Explainable AI* (LIME, SHAP) para fortalecer la transparencia y confianza regulatoria del sistema.

Financiamiento: Esta investigación no recibió financiamiento externo.

Referencias

Adnan, M., Xiao, B., Ali, M. U., Bibi, S., Yu, H., Xiao, P., Zhao, P., Wang, H., & An, X. (2024). Human Inventions and its Environmental Challenges, especially Artificial Intelligence: New Challenges Require New Thinking. *Environmental Challenges*.
<https://doi.org/10.1016/j.envc.2024.100976>

Aletras, N., Tsarapatsanis, D., Preoțiu-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A Natural Language Processing perspective. *PeerJ Computer Science*, 2, e93. <https://doi.org/10.7717/peerj-cs.93>

Aman, A., & Reji, D. J. (2022). EnvBert: An NLP model for Environmental Due Diligence data classification. *Software Impacts*, 14, 100427. <https://doi.org/10.1016/j.simpa.2022.100427>

Anh, D. H., Dinh-Truong, Tran, V., & Minh, N. L. (2023). The Impact of Large Language Modeling on Natural Language Processing in Legal Texts: A Comprehensive Survey. 2023 15th International Conference on Knowledge and Systems Engineering (KSE), 1-7.



<https://doi.org/10.1109/kse59128.2023.10299488>

Ariai, F., Mackenzie, J., & Demartini, G. (2025). Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges (arXiv:2410.21306). arXiv.

<https://doi.org/10.48550/arXiv.2410.21306>

Athukorala, S., & Mohotti, W. (2022). An effective short-text topic modelling with neighbourhood assistance-driven NMF in Twitter. *Social Network Analysis and Mining*, 12(1), 89. <https://doi.org/10.1007/s13278-022-00898-5>

Beatrice Oyinkansola Adelakun, Bernard Owusu Antwi, Afari Ntiakoh, & Augustine Obinna Eziefule. (2024). Leveraging AI for sustainable accounting: Developing models for environmental impact assessment and reporting. *Finance & Accounting Research Journal*, 6(6), 1017-1048. <https://doi.org/10.51594/farj.v6i6.1234>

Bond, A., Cilliers, D., Retief, F., Alberts, R., Roos, C., & Moolman, J. (2024). Using an Artificial intelligence chatbot to critically review the scientific literature on the use of Artificial intelligence in Environmental Impact Assessment. *Impact Assessment and Project Appraisal*, 42(2), 189-199. <https://doi.org/10.1080/14615517.2024.2320591>

Campillos-Llanos, L., Valverde-Mateos, A., & Capllonch-Carrión, A. (2025). Hybrid natural language processing tool for semantic annotation of medical texts in Spanish. *BMC Bioinformatics*, 26. <https://doi.org/10.1186/s12859-024-05949-6>

Cuevas, J. P., Reyes-Ortíz, J., Cuevas-Rasgado, A., Mora-Gutiérrez, R.-A., & Bravo, M. (2024). MédicoBERT: A Medical Language Model for Spanish Natural Language Processing Tasks with a Question-Answering Application Using Hyperparameter Optimization. *Applied*

Sciences.

<https://doi.org/10.3390/app14167031>

Garmendia Salvador, A., Salvador Alcaide, A., Crespo Sánchez, C., & Garmendia Salvador, L. (2005). Evaluación de impacto ambiental. PEARSON EDUCACIÓN, S.A.

Haq, H. U., Kocaman, V., & Talby, D. (2022). Deeper Clinical Document Understanding Using Relation Extraction.

Holtgrieve, D. G., & Sanford, R. M. (2022). Environmental Impact Assessment in the United States (1.a ed.). Routledge. <https://doi.org/10.4324/9781003030713>

Houser, H. (2020). Infowhelm: Environmental art and literature in an age of data. Columbia University Press.

Islam, M. N., Hossain, M. M., & Ornob, M. S. S. (2024). Business research on Industry 4.0: A systematic review using topic modelling approach. *Future Business Journal*, 10(1), Article 1. <https://doi.org/10.1186/s43093-024-00398-2>

Jagannathan, M., Roy, D., & Delhi, V. S. K. (2022). Application of NLP-based topic modeling to analyse unstructured text data in annual reports of construction contracting companies. *CSI Transactions on ICT*, 10(2), Article 2. <https://doi.org/10.1007/s40012-022-00355-w>

Leal, A., Maldonado, S., Martínez, J. I., Bertazzo, S., Quijada, S., & Vairetti, C. (2025). Enhancing environmental governance: A text-based artificial intelligence approach for project evaluation involvement. *Environmental Impact Assessment Review*. <https://doi.org/10.1016/j.eiar.2024.107707>

Lee, D., & Seung, H. S. (2000). Algorithms for Non-negative Matrix Factorization. *Advances in Neural Information Processing Systems*, 13. https://proceedings.neurips.cc/paper_files/paper/2000/hash/f9d1152547c0bde01830b7e8bd60024c-Abstract.html



- Licari, D., & Comandé, G. (2024). ITALIAN-LEGAL-BERT models for improving natural language processing tasks in the Italian legal domain. *Comput. Law Secur. Rev.*, 52, 105908. <https://doi.org/10.1016/j.clsr.2023.105908>
- Morrison-Saunders, A. (2023). *Advanced introduction to environmental impact assessment (Second edition)*. Edward Elgar Publishing.
- Naciones Unidas. (2025). *The-Sustainable-Development-Goals-Report-2025*. <https://unstats.un.org/sdgs/report/2025/The-Sustainable-Development-Goals-Report-2025.pdf>
- Pacto Mundial. (2026). [Info.pactomundial.org/1/867062/2026-01-30/t5nj8l/867062/1769768600cfYf29nZ/Publicaci_n_Pacto_Mundial__IA_y_Empresas.pdf](https://info.pactomundial.org/1/867062/2026-01-30/t5nj8l/867062/1769768600cfYf29nZ/Publicaci_n_Pacto_Mundial__IA_y_Empresas.pdf). Pacto Mundial de la ONU España. https://info.pactomundial.org/1/867062/2026-01-30/t5nj8l/867062/1769768600cfYf29nZ/Publicaci_n_Pacto_Mundial__IA_y_Empresas.pdf
- Pauzi, Z., & Capiluppi, A. (2023). Applications of natural language processing in software traceability: A systematic mapping study. *Journal of Systems and Software*, 198, 111616. <https://doi.org/10.1016/j.jss.2023.111616>
- Princy, Rufina., Gawali, P., Murari, K., Agrawal, T., Kavitha, R., & Mukuntharaj, C. (2023). Natural Language Processing for Extracting Meaningful Insights from Textual Information. 2023 2nd International Conference on Futuristic Technologies (INCOFT), 1-5. <https://doi.org/10.1109/incoft60753.2023.10425508>
- Sanchez, L. E. (2008). *Origen y Difusión de la Evaluación de Impacto Ambiental. En Evaluación del Impacto Ambiental—Conceptos y Métodos*. Ecoe Ediciones.
- Silveira, R., Ponte, C., Almeida, V., Pinheiro, V., & Furtado, V. (2023). LegalBert-pt: A Pretrained Language Model for the Brazilian Portuguese Legal Domain. 268-282. https://doi.org/10.1007/978-3-031-45392-2_18
- Spasic, I., & Nenadic, G. (2020). Clinical Text Data in Machine Learning: Systematic Review. *JMIR Medical Informatics*, 8(3), e17984. <https://doi.org/10.2196/17984>
- Tironi, M., & Lisboa, D. I. R. (2023). Artificial intelligence in the new forms of environmental governance in the Chilean State: Towards an eco-algorithmic governance. *Technology in Society*. <https://doi.org/10.1016/j.techsoc.2023.102264>
- Zhang, L., Suo, W., Zhang, Z., & Liu, M. (2022). Extending Bayesian Learning to Intelligent Learning in Human Way. *Procedia Computer Science*, 199, 118-124. <https://doi.org/10.1016/j.procs.2022.01.015>