

Explicaciones lingüísticas difusas enriquecidas para la predicción de fuga y segmentación de clientes en telecomunicaciones: extensión y aplicación de SHAP-Rule

Enhanced Fuzzy Linguistic Explanations for Customer Churn Prediction and Customer Segmentation in Telecommunications: An Extension of SHAP-Rule

Eduardo Espinoza, Elvis Ramirez, Lesli Diaz, Rodeney Llantoy, Sayumi Canchari

Maestría en Inteligencia Artificial

Universidad Nacional Mayor de San Marcos

RESUMEN

Las explicaciones basadas en valores de Shapley permiten identificar la contribución de cada variable a una predicción, pero su representación numérica resulta difícil de interpretar para usuarios no especializados. Por otro lado, las reglas lingüísticas difusas mejoran la legibilidad, aunque pierden el detalle cuantitativo de las atribuciones. Este trabajo propone la Explicación Difusa Enriquecida, una capa adicional que complementa las reglas generadas por SHAP-R con información estadística y de importancia local, sin modificar la regla original ni recalcular las atribuciones. La propuesta se evalúa utilizando un conjunto público de 7 043 clientes de telecomunicaciones en dos tareas: predicción de fuga, mediante modelos de Random Forest y LightGBM, y segmentación de clientes obtenida con análisis de componentes principales y K-Means, posteriormente explicada mediante un clasificador auxiliar Random Forest. Los resultados muestran que las reglas lingüísticas reducen la complejidad de las explicaciones de diez condiciones a un promedio de entre 3,1 y 3,6, al tiempo que incrementan la consistencia textual de 0,60 a 0,81, manteniendo coberturas entre 0,90 y 0,98. En la tarea de segmentación, las explicaciones alcanzan una consistencia de 0,97, permitiendo describir los segmentos en términos de variables del negocio. Además, el estudio identifica una limitación no reportada previamente en SHAP-R: las variables con alta concentración de valores idénticos generan particiones lingüísticas degeneradas que reducen la capacidad discriminativa de las reglas.

Palabras clave: inteligencia artificial explicable; valores de Shapley; reglas difusas; variables lingüísticas; fuga de clientes; segmentación de clientes; telecomunicaciones; sistemas de apoyo a la decisión.

I. INTRODUCCIÓN

La fuga de clientes es un gran desafío para la industria de las telecomunicaciones, perder abonados reduce los ingresos recurrentes e incrementa los costos de adquisición de nuevos clientes. Mettle et al. (2026) demostraron que la aplicación de modelos de aprendizaje automático permite anticipar con alta precisión la probabilidad de deserción, con información que favorece a retener clientes valiosos, optimizar la asignación de recursos comerciales y mejorar el desempeño financiero.

La inteligencia artificial ya no es una herramienta para incrementar la exactitud predictiva, sino un soporte para la toma de decisiones que transforma los modelos de predicción en instrumentos de generación de valor mediante el entendimiento de los clientes y su recuperación y conservación.

La interpretabilidad es un requisito esencial para la adopción de los modelos de inteligencia artificial. Los modelos ofrecen un elevado desempeño predictivo pero su naturaleza de «caja negra» dificulta la comprensión del sustento de sus predicciones. El estudio sobre inteligencia artificial explicable de Sahakyan et al. (2021) señala que métodos como SHAP se han consolidado como el estándar para explicar modelos complejos debido a su sólido fundamento teórico y su capacidad para cuantificar la contribución de cada variable. Pero también destaca que las explicaciones continúan presentándose principalmente mediante valores numéricos y representaciones gráficas que requieren conocimientos especializados para su correcta interpretación.

Esto evidencia la necesidad de desarrollar mecanismos que traduzcan las explicaciones en información más comprensible y accionable para los usuarios del negocio.

II. OBJETIVOS

A. Objetivo general

Diseñar y evaluar una extensión enriquecida de las explicaciones lingüísticas difusas derivadas de valores de Shapley, que incorpore a cada antecedente su contexto estadístico y validar su aplicabilidad en dos estudios del sector de telecomunicaciones: la predicción de fuga y la asignación de segmentos.

B. Objetivos específicos

Diseñar e implementar reglas de explicación lingüística difusa generadas a partir del algoritmo SHAP-R, enriqueciendo cada explicación con contexto estadístico.

Aplicar las explicaciones lingüísticas difusas enriquecidas a modelos de clasificación de fuga de clientes y determinar en qué medida las reglas resultantes identifican y jerarquizan los factores que gobiernan cada predicción individual.

Aplicar las explicaciones lingüísticas difusas enriquecidas a un clasificador de segmentos obtenidos mediante reducción de dimensionalidad y agrupamiento, y determinar si permiten expresar la pertenencia de un cliente a un segmento en términos de las variables originales del negocio.

III. FUNDAMENTOS

A. Antecedentes

1) Modelos de machine learning

Cyril (2026) propuso un modelo híbrido de ensamble que integra LightGBM y Random Forest con el objetivo de incrementar la precisión de predicción de clientes propensos a abandonar el servicio de telecomunicaciones. Reportó mejoras respecto a modelos individuales, mostrando que la combinación de algoritmos puede capturar patrones de comportamiento más complejos presentes en los datos.

Ali et al. (2026) propusieron un marco de optimización de marketing para telecomunicaciones que integra predicción de churn, segmentación de clientes mediante K-Means y estrategias de retención específicas para cada segmento. Utilizando el conjunto de datos IBM Telco Customer Churn, determinaron el número de clústeres mediante el método Elbow y los visualizaron con PCA, obteniendo segmentos de clientes de alto, medio y bajo valor. Posteriormente, combinaron estos segmentos con el riesgo de abandono para definir estrategias diferenciadas de retención y explicaron las predicciones mediante SHAP, demostrando que la integración de segmentación y modelos predictivos produce decisiones de marketing más accionables que la predicción de churn por sí sola.

2) Interpretabilidad con SHAP

Zerine et al. (2026) presentaron un estudio comparativo de cinco modelos de aprendizaje automático para telecomunicaciones, complementando la evaluación predictiva con un análisis de interpretabilidad mediante SHAP. El trabajo permitió identificar las variables con mayor influencia sobre la deserción, destacando factores como el tipo de contrato, los cargos mensuales y la antigüedad del cliente.

Antoni et al. (2025) desarrollaron un modelo de aprendizaje, para telecomunicaciones, por apilamiento

optimizado mediante un meta-aprendiz ajustado, complementado con explicaciones generadas por SHAP. El estudio obtuvo elevados niveles de precisión y determinó que variables como el número de referidos, el tipo de contrato, el cargo mensual y la permanencia del cliente representan los principales factores asociados al abandono.

3) Explicaciones lingüísticas difusas

Khalyasmaa et al. (2025) propusieron SHAP-Rule, un método de inteligencia artificial explicable orientado a transformar los valores de contribución generados por SHAP en reglas lingüísticas difusas del tipo SI-ENTONCES. La metodología combina la solidez teórica de los valores de Shapley con la lógica difusa para generar explicaciones más comprensibles para usuarios no especializados, incorporando además una medida cuantitativa de la fuerza de activación de cada regla. Los autores validaron el método en problemas de diagnóstico de rodamientos industriales, transformadores de potencia y diabetes, demostrando que las reglas generadas conservan la capacidad explicativa de SHAP y mejoran significativamente la interpretabilidad de los modelos de aprendizaje automático.

B. Marco teórico

1) Random Forest

Random Forest es un ensemble de árboles de decisión que agrega sus predicciones por voto mayoritario, combinando remuestreo bootstrap de instancias con selección aleatoria de características en cada nodo. Esta doble aleatorización mantiene los árboles débilmente correlacionados y reduce la varianza del predictor agregado, mecanismo que puede interpretarse como una regularización implícita (Scornet & Hooker, 2026). Los ensembles de árboles siguen siendo el estado del arte en datos tabulares de escala media frente a arquitecturas profundas, por su robustez ante características no informativas y su capacidad de aproximar funciones irregulares (Grinsztajn et al., 2022). Al ser un modelo de árboles, admite atribuciones exactas mediante TreeSHAP (Chen et al., 2023).

2) LightGBM

LightGBM implementa *gradient boosting* de árboles, construyendo estimadores secuencialmente sobre el gradiente residual mediante una aproximación de Taylor de segundo orden de la pérdida regularizada. Tres mecanismos lo distinguen de otras implementaciones de GBDT (So, 2024): **GOSS**, que conserva las instancias de gradiente elevado y submuestra las restantes con un factor de compensación; **EFB**, que agrupa variables mutuamente excluyentes en espacios dispersos; y el **crecimiento leaf-wise**, que expande la hoja de mayor reducción de pérdida en lugar de crecer por niveles, alcanzando menor error con igual número de nodos a costa de un mayor riesgo de sobreajuste, mitigable limitando la profundidad. Junto con la discretización por histogramas, estas optimizaciones reducen el coste de búsqueda de cortes, manteniendo la compatibilidad con TreeSHAP (Chen et al., 2023).

3) Inteligencia artificial explicable

La XAI agrupa los métodos que hacen comprensible el comportamiento de un modelo para una audiencia determinada (Ali et al., 2023), y resulta crítica en diagnóstico industrial y médico, donde las decisiones deben justificarse y comunicarse con rapidez (Alexander et al., 2024; Cacao et al., 2025). Se clasifican según tres ejes: transparencia intrínseca frente a explicabilidad *post hoc*, alcance local frente a global, y métodos agnósticos al modelo frente a específicos (Ali et al., 2023). Su evaluación requiere métricas complementarias: **fidelidad** (acuerdo con el modelo original), **cobertura** (proporción de instancias explicadas sobre un umbral de calidad), **complejidad** (número de condiciones, asociada a la carga cognitiva), **consistencia** (estabilidad ante perturbaciones de la entrada) y **valoración experta**, dado que las métricas cuantitativas no determinan por sí solas la utilidad práctica (Cao et al., 2024). Un hallazgo transversal es que las atribuciones numéricas resultan cognitivamente costosas cuando se exige comprensión inmediata, frente a las representaciones simbólicas SI-ENTONCES, especialmente las difusas (Cao et al., 2024; Khalyasmaa et al., 2025; Ouifak & Idri, 2025).

4) Valores de Shapley

Los valores de Shapley reparten el valor de un juego cooperativo promediando la contribución marginal de cada jugador sobre todos los órdenes de formación de coaliciones, como expresa (1).

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(n-|S|-1)!}{n!} [v(S \cup \{j\}) - v(S)] \quad (1)$$

Es la única asignación que satisface eficiencia, simetría, jugador nulo y aditividad. Aplicados a la atribución de características —variables como jugadores, predicción como valor del juego— dan lugar al modelo aditivo $g(z') = \phi_0 + \sum_j \phi_j z'_j$, con ϕ_0 el valor esperado de la predicción (Chen et al., 2023). Su estimación combina dos decisiones: cómo se elimina la información de una característica y qué estrategia tractable se emplea, ya que el cálculo exacto es exponencial (Chen et al., 2023). De ahí derivan aproximaciones agnósticas (KernelSHAP, muestreo de permutaciones) y específicas, entre las que **TreeSHAP** calcula valores exactos en tiempo polinómico $O(T \cdot L \cdot D^2)$ para ensembles de árboles, viabilizando su uso en tiempo real. Cabe advertir que distintas convenciones de eliminación producen atribuciones no intercambiables, por lo que debe explicitarse la variante empleada (Chen et al., 2023; Verdinelli & Wasserman, 2024).

5) SHAP-Rule

SHAP-Rule convierte atribuciones SHAP en reglas SI-ENTONCES difusas con activación cuantificada, actuando como capa de explicación *post hoc* de un solo nivel sobre un clasificador entrenado: emite una regla local por instancia, sin base global ni inferencia jerárquica (Khalyasmaa et al., 2025). El procedimiento comprende cinco etapas. (i) Los siete términos lingüísticos, de «extremadamente bajo» a «extremadamente alto», se derivan automáticamente calculando las fronteras sobre escala $L = \text{arcsinh}(x)$ —que admite valores negativos y atenúa la asimetría— con el criterio de Tukey de $1,5 \cdot \text{IQR}$, retransformando mediante $\sinh$ y podando términos con fronteras coincidentes. (ii) Sobre cada terna consecutiva se construyen funciones de pertenencia triangulares. (iii)

Las atribuciones se normalizan, $w_j = \phi_j / \sum_i |\phi_i|$. (iv) Ordenadas por $|w_j|$, se retienen $q = \min(m, K)$ características, donde m es el menor índice cuya suma acumulada alcanza el umbral T ; este doble criterio adapta el número de condiciones a la concentración real de la importancia. (v) La regla se ensambla y su intensidad se calcula como media geométrica ponderada, según (2).

$$\alpha = \prod_{j=1}^q \mu_j(x_j)^{1-|w_j|} \quad (2)$$

La expresión integra importancia y grado de ajuste al término. La complejidad adicional es $O(F \log F)$ por instancia.

6) SHAP-L

SHAP-L es la explicación numérica de TreeExplainer restringida a las características más influyentes mediante el mismo criterio de selección que SHAP-Rule (umbral acumulado T y tope K), diferenciándose solo en el formato de salida: conserva valores y atribuciones numéricas en lugar de términos lingüísticos (Khalyasmaa et al., 2025).

IV. METODOLOGÍA

La metodología se organiza en cuatro etapas encadenadas: adquisición de datos, análisis exploratorio y preparación, modelado con aprendizaje automático y explicabilidad mediante SHAP-Rule Enhanced. Las dos últimas se desdoblan en dos estudios que comparten el mismo sustrato de datos, pero difieren en la variable objetivo: el Estudio 1 predice la fuga de clientes y el Estudio 2 predice el segmento al que pertenece cada cliente, descubierto previamente mediante agrupamiento no supervisado.

A. Adquisición de datos

El estudio emplea el conjunto **Telecom Customer Churn** de Maven Analytics (2022), que registra 7 043 clientes de una compañía ficticia de telecomunicaciones de California correspondientes al segundo trimestre de 2022, descritos por 38 atributos de perfil demográfico, ubicación, antigüedad, servicios contratados, facturación y estado final. Distingue el estado del cliente en tres categorías en lugar de un indicador binario; y ofrece un mayor número de variables numéricas, condición necesaria para un método que genera antecedentes difusos sobre atributos continuos.

B. Análisis exploratorio y preparación de los datos

División del conjunto. La partición precede a la limpieza. Es estratificada por el objetivo, en proporción 80/20 y con semilla fija, lo que garantiza que ninguna operación posterior aprenda parámetros de las observaciones de prueba. Ambos estudios comparten la misma partición, de modo que sus resultados son directamente comparables.

Limpieza. Se aplican dos correcciones. La primera es una imputación estructural basada en reglas del diccionario de datos y no en estadísticos: nueve servicios dependientes de conexión a internet quedan vacíos por diseño en los clientes que no la contratan y se codifican con una categoría explícita, mientras que la ausencia de servicio de larga distancia o de descarga de datos se

traduce en un valor real de cero. Al no aprender parámetro alguno, se aplica de forma idéntica a ambas particiones. La segunda sustituye los valores negativos del cargo mensual, créditos y reembolsos se registran aparte, por la mediana de los valores válidos de entrenamiento.

Análisis exploratorio. Sobre la partición de entrenamiento ya limpia se examinan distribuciones condicionadas al objetivo, asimetría, concentración en valores modales y correlaciones. El análisis cumple una función instrumental: la asimetría de las variables monetarias justifica la transformación empleada en la capa explicativa, y la concentración en cero de varias variables anticipa el comportamiento degenerado documentado en los resultados. La Tabla 1 resume los estadísticos descriptivos obtenidos.

Selección de variables. El modelado se restringe a las variables numéricas, ya que el algoritmo SHAP-R está diseñado para este tipo de variables. Sobre ellas se elimina la multicolinealidad mediante el factor de inflación de la varianza, aplicado iterativamente hasta que todas queden por debajo del umbral de 10, lo que deja 10 variables finales. El criterio responde a un requisito de la capa explicativa: cuando dos variables son casi redundantes, el valor de Shapley reparte la contribución entre ambas y diluye la atribución individual, de modo que variables relevantes pueden quedar fuera de la regla por fragmentación de su importancia. Este conjunto de diez variables constituye el espacio de entrada de ambos estudios en el aprendizaje supervisado.

Tabla 1. Estadísticos descriptivos de las variables numéricas sobre la partición de entrenamiento y proporción de la moda. Fuente: elaboración propia.

Variable	Media	Mediana	Asimetría	Prop. moda
Edad	46,76	46,00	0,16	0,023
Número de Dependientes	0,47	0,00	2,10	0,769
Número de Referidos	2,02	0,00	1,39	0,524
Antigüedad en Meses	34,63	33,00	0,14	0,062
Cargo Prom. Mensual de Larga Distancia	23,06	22,95	0,05	0,096
Promedio Mensual de GB Descargados	20,84	17,00	1,21	0,205
Cargo Mensual	66,43	73,15	-0,30	0,017
Reembolsos Totales	2,14	0,00	4,13	0,918
Cargos Totales por Datos Extra	7,12	0,00	4,01	0,899
Cargos Totales de Larga Distancia	803,02	476,14	1,14	0,096

C. Modelado con aprendizaje automático

1) Estudio 1: predicción del churn

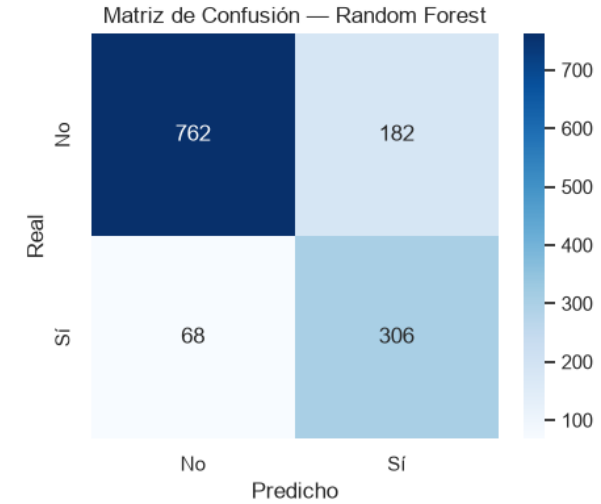
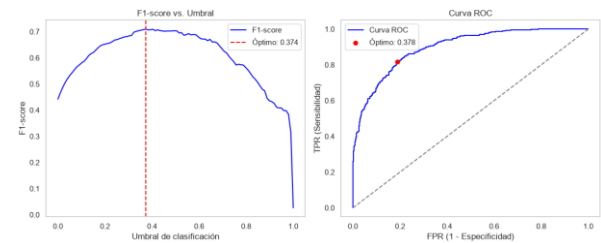
Se entrenan dos ensambles de árboles, **Random Forest** y **LightGBM**, según propone Cyril (2026) y la disponibilidad de TreeSHAP como algoritmo de atribución exacto.

El ajuste se realiza con Optuna, evaluando 30 configuraciones por modelo bajo validación cruzada estratificada de cinco pliegues y optimizando el F1 promedio; los espacios de búsqueda se detallan en la Tabla 2. El desbalance se compensa con SMOTE.

El punto de corte del entrenamiento no se fija en 0,5: se barren cien umbrales sobre las probabilidades fuera de pliegue de entrenamiento y se selecciona el que maximiza el F1, fijándolo antes de evaluar sobre prueba. La decisión condiciona la capa explicativa, ya que el consecuente de cada regla debe corresponder a la predicción efectiva del modelo bajo su umbral operativo. El desempeño se evalúa con exactitud, precisión, sensibilidad, F1, AUC-ROC y matriz de confusión (Figuras 1 y 2).

Tabla 2. Espacios de búsqueda de hiperparámetros. Fuente: elaboración propia.

Modelo	Hiperparámetro	Rango
Random Forest	Número de árboles	50–300
	Profundidad máxima	3–20
	Mínimo de muestras para dividir / por hoja	2–10 / 1–10
	Selección de variables por división	sqrt, log2
LightGBM	Número de hojas	7–127
	Profundidad máxima	3–12
	Tasa de aprendizaje	0,01–0,3 (log)
	Mínimo de muestras por hoja	5–100
	Submuestreo de filas y columnas	0,5–1,0
	Regularización L1 y L2	1×10^{-8} –10 (log)



```

RF -> umbral óptimo (F1): 0.374
===== Random Forest =====

```

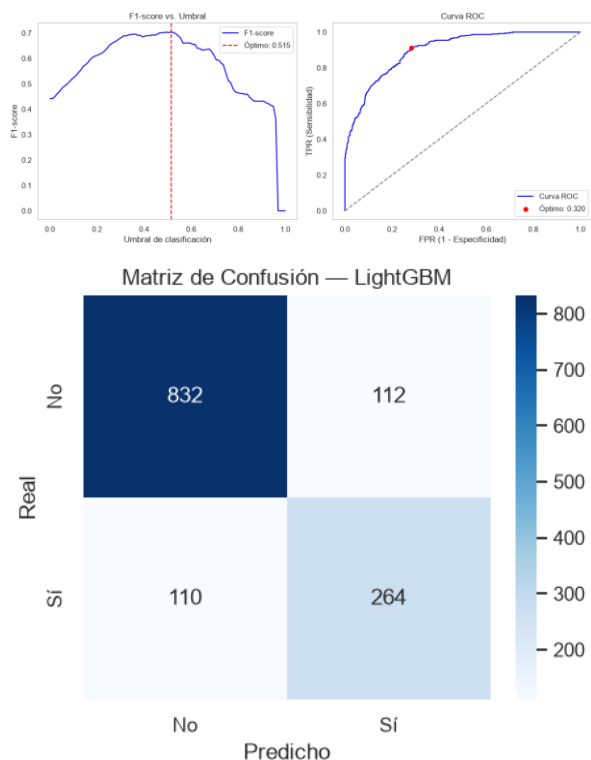
	precision	recall	f1-score	support
No churn	0.92	0.81	0.86	944
Churn	0.63	0.82	0.71	374
accuracy			0.81	1318
macro avg	0.77	0.81	0.78	1318
weighted avg	0.84	0.81	0.82	1318

```

Accuracy: 0.81
Precision: 0.63
Recall: 0.82
F1-score: 0.71
AUC-ROC: 0.81

```

Figura 1. Métricas de desempeño, curva ROC y matriz de confusión del modelo Random Forest en la predicción de fuga. Fuente: elaboración propia.



```

LightGBM -> umbral óptimo (F1): 0.515
===== LightGBM =====

```

	precision	recall	f1-score	support
No churn	0.88	0.88	0.88	944
Churn	0.70	0.71	0.70	374
accuracy			0.83	1318
macro avg	0.79	0.79	0.79	1318
weighted avg	0.83	0.83	0.83	1318

```

Accuracy: 0.83
Precision: 0.70
Recall: 0.71
F1-score: 0.70
AUC-ROC: 0.79

```

Figura 2. Métricas de desempeño, curva ROC y matriz de confusión del modelo LightGBM en la predicción de fuga. Fuente: elaboración propia.

2) Estudio 2: descubrimiento y predicción de segmentos

El segundo estudio combina aprendizaje no supervisado y supervisado sobre la misma población y

partición. Su objetivo no es predecir un desenlace observado, sino descubrir estructura latente en la base de clientes y construir después un clasificador capaz de asignar clientes al segmento correspondiente, de modo que la capa explicativa pueda enunciar en términos lingüísticos por qué un cliente pertenece a un segmento.

Preprocesamiento. El agrupamiento requiere incorporar las variables categóricas, ya que el perfil contractual y de servicios contribuye a definir los segmentos. Se aplica codificación one-hot a las 18 categóricas y estandarización a las 10 numéricas retenidas tras el filtrado por factor de inflación de la varianza, lo que produce un espacio de 58 columnas.

Reducción de dimensionalidad. Se aplica análisis de componentes principales conservando 16 componentes, que acumulan el 86,1 % de la varianza (Figura 3). La proyección mejora de forma consistente la estructura del agrupamiento: el coeficiente de silueta aumenta en todos los valores de k evaluados, y en particular pasa de 0,183 a 0,215 para $k = 4$ y de 0,212 a 0,252 para $k = 2$.

Selección del número de segmentos. Se evaluaron valores de k entre 2 y 10 mediante el método del codo y el coeficiente de silueta (Figura 4). El máximo de silueta corresponde a $k = 2$ (0,252 en el espacio proyectado), pero se adopta $k = 4$ por dos razones (Figura 5). Primera, el análisis de anidamiento muestra que la partición en cuatro segmentos es un refinamiento estrictamente contenido en la de dos, con pureza igual a 1,0 en los cuatro subgrupos, de modo que $k = 4$ no contradice la estructura dominante sino que la desagrega. Segunda, la partición en dos segmentos separa esencialmente a los clientes con y sin servicio de internet, una distinción de escaso valor operativo para un equipo de retención, mientras que $k = 4$ distingue perfiles con tasas de fuga que van del 9 % al 49 %. La Figura 6 presenta la silueta por cliente para la solución adoptada.

Agrupamiento. Sobre el espacio proyectado se aplica k -means con $k = 4$, obteniendo cuatro segmentos etiquetados de 0 a 3. El codificador, el estandarizador, la proyección y los centroides se ajustan exclusivamente sobre entrenamiento; las observaciones de prueba se proyectan y se asignan al centroide más próximo sin reajustar el modelo, evitando que la definición de los segmentos incorpore información del conjunto de evaluación.

Clasificación de segmentos. Las etiquetas de agrupamiento se emplean como variable objetivo de un clasificador Random Forest entrenado únicamente sobre las diez variables numéricas en su escala original, no sobre las componentes principales. Esta restricción es deliberada y constituye el eje del diseño: los antecedentes de las reglas deben expresarse en unidades interpretables por el usuario, y una condición del tipo «la segunda componente principal es alta» carecería de significado operativo. El modelo actúa por tanto como sustituto del agrupamiento, y la validez de las explicaciones depende de su fidelidad respecto de las asignaciones reales de k -means, que se reporta en la Figura 7. Los hiperparámetros se ajustan mediante validación cruzada estratificada de cinco pliegues optimizando el F1 macropromediado.

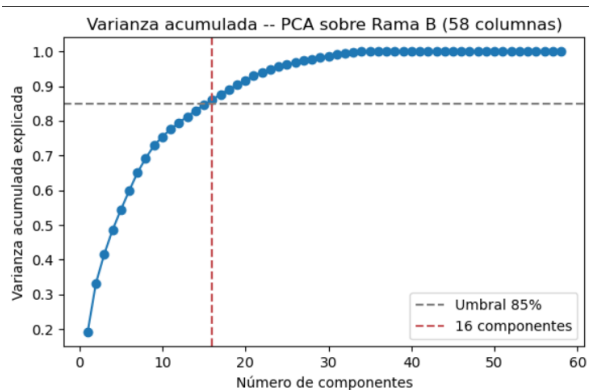


Figura 3. Proyección de los clientes mediante análisis de componentes principales. Fuente: elaboración propia.

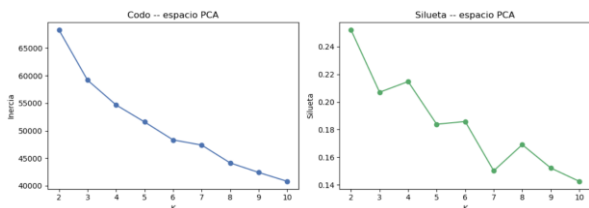


Figura 4. Método del codo y coeficiente de silueta para la selección del número de segmentos. Fuente: elaboración propia.

=== COMPARACIÓN: silueta original (58 cols) vs. silueta con PCA ===

K	Silueta original	Silueta PCA
2	0.212	0.252 <-- K=2 (óptimo sin PCA)
3	0.178	0.207
4	0.183	0.215 <-- K=4
5	0.166	0.184
6	0.155	0.186
7	0.143	0.150
8	0.143	0.169
9	0.146	0.152
10	0.116	0.143

K óptimo por silueta en espacio PCA : 2 (silueta = 0.252)

K=4 -- silueta original: 0.183 | silueta con PCA: 0.215 (cambio: +0.032)

Figura 5. Justificación de la elección de k = 4: anidamiento de las particiones y tasa de fuga por segmento. Fuente: elaboración propia.

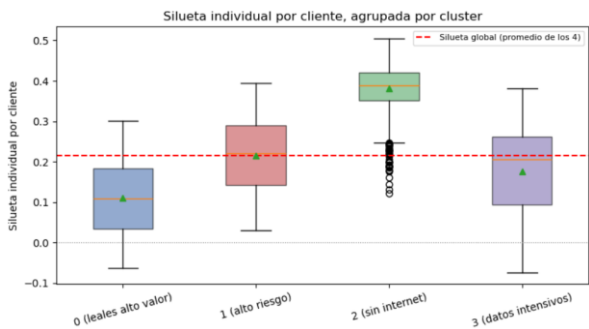


Figura 6. Coeficiente de silueta por cliente para la partición en cuatro segmentos. Fuente: elaboración propia.

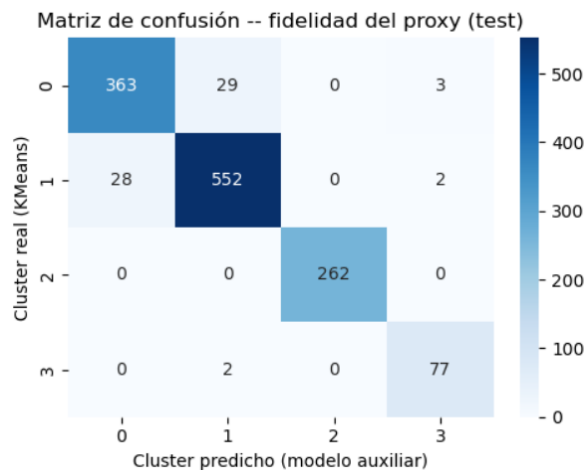


Figura 7. Matriz de confusión y métricas del clasificador Random Forest sustituto del agrupamiento. Fuente: elaboración propia.

D. Generación de reglas lingüísticas difusas

1) SHAP-Rule Enhanced

Las atribuciones se calculan con **TreeSHAP** y se traducen mediante **SHAP-Rule**, una capa post hoc de un solo nivel que emite una regla local por instancia. El procedimiento consta de cuatro pasos.

Primero, para cada variable se derivan siete términos lingüísticos —de «extremadamente bajo» a «extremadamente alto»— calculando las fronteras sobre escala arcsinh, que atenúa la asimetría y admite valores negativos, con el criterio de Tukey de $1.5 \cdot \text{IQR}$ y poda de términos redundantes.

Segundo, sobre cada terna consecutiva se construyen funciones de pertenencia triangulares.

Tercero, las atribuciones se normalizan y se seleccionan los antecedentes mediante un doble criterio de umbral acumulado T y tope máximo K , que adapta la longitud de la regla a la concentración real de la importancia.

Cuarto, la regla se ensambla como conjunción de pares variable-término y su validez se cuantifica mediante una media geométrica ponderada de los grados de pertenencia. Los parámetros se fijan en los valores de la formulación original ($K = 4$, $T = 0.80$, $\alpha \geq 0.33$) para permitir la comparación directa, y las fronteras lingüísticas, derivadas de entrenamiento, son idénticas en ambos estudios.

2) Capa de enriquecimiento

La regla difusa condensa la predicción en etiquetas lingüísticas y, al hacerlo, descarta la información cuantitativa subyacente: no distingue entre un cliente de un mes y uno de once dentro del mismo término, no indica el peso relativo de cada antecedente ni la proximidad a la frontera con el término contiguo. Como aporte de este

trabajo se propone una capa de extensión que, para cada antecedente, construye una estructura acompañada de dos visualizaciones deterministas: el panel de la función de pertenencia con la posición del cliente y el histograma de la distribución de la variable en entrenamiento. La capa es estrictamente aditiva: consume las salidas de la etapa anterior sin recalculat atribuciones ni alterar la regla, por lo que no puede divergir de la explicación original ni añade coste sobre la ruta crítica de generación.

3) Aplicación a los dos estudios

En el **Estudio 1**, la explicación enriquecida responde a por qué un cliente concreto es clasificado como riesgo de fuga. En el **Estudio 2**, responde a por qué un cliente pertenece a un segmento determinado, lo que convierte una partición estadística en una caracterización comunicable: en lugar de describir el segmento por sus centroides en el espacio de componentes principales, la regla lo describe por los valores lingüísticos de las variables originales que determinan la pertenencia.

E. Métricas de evaluación de las explicaciones

Se comparan tres formatos sobre las mismas instancias y bajo todos los modelos: **SHAP**, que enumera todas las variables con su valor y atribución; **SHAP-L**, que aplica el mismo criterio de selección, pero conserva el registro

Tabla 3. Términos lingüísticos generados sobre la partición de entrenamiento. Ninguna variable generó el término «extremadamente alto». Fuente: elaboración propia.

Variable	N.º de términos	Extrem. bajo	Bajo	Por debajo del prom.	Promedio	Por encima del prom.	Alto
Edad (años)	5	—	19	33	46	60	80
Número de Dependientes	1	—	—	—	0	—	—
Número de Referidos	3	—	—	—	0	3	11
Antigüedad en Meses	5	—	1	12	33	57	72
Cargo Prom. Mensual Larga Distancia (USD)	6	0	1,00	9,29	22,95	36,46	49,99
Promedio Mensual de GB Descargados	5	—	0	4	17	27	85
Cargo Mensual (USD)	5	—	18,25	41,15	73,15	90,37	118,75
Reembolsos Totales (USD)	1	—	—	—	0	—	—
Cargos Totales por Datos Extra (USD)	1	—	—	—	0	—	—
Cargos Totales Larga Distancia (USD)	6	0	2,43	106,54	476,14	1 289,73	3 564,72

En la Figura 8 se pueden observar los gráficos de las funciones de pertenencia de las variables que obtuvieron más de un término.

numérico; y **SHAP-R**, la regla lingüística difusa. SHAP-L opera como control experimental: al compartir con SHAP-R el criterio exacto de selección y diferir solo en el formato de salida, permite aislar el efecto de la verbalización lingüística del de la reducción de complejidad.

La evaluación emplea tres métricas:

Complejidad: número de condiciones de la explicación.

Consistencia: similitud de Levenshtein normalizada entre el texto original y el generado para diez perturbaciones gaussianas de la instancia, con desviación equivalente al 1 % de la desviación estándar de cada variable.

Cobertura: proporción de instancias cuya regla alcanza la activación mínima.

V. RESULTADOS

A. Términos lingüísticos generados

Los términos lingüísticos generados para las 10 variables numéricas se observan en la Tabla 3, en donde también se observa que 3 variables tienen un solo término debido a la alta concentración de ceros en los valores de las variables.

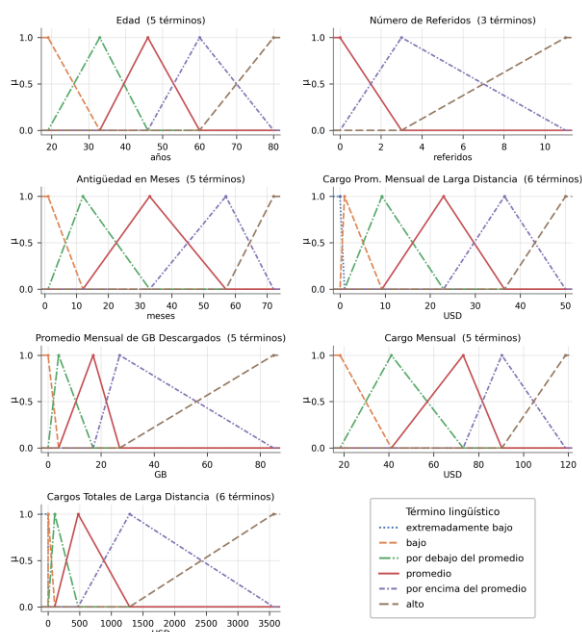


Figura 8. Funciones de pertenencia triangulares de las variables con más de un término lingüístico. Fuente: elaboración propia.

B. Explicaciones generadas

1) Estudio 1: predicción del churn

En la Tabla 4 se muestra la comparación entre el resultado numérico de SHAP y SHAP-L y el resultado explicativo con lenguaje difuso de SHAP-R.

Tabla 4. LightGBM: comparativa de los formatos explicativos SHAP, SHAP-L y SHAP-R. Fuente: elaboración propia.

Instancia	Formato	Explicación generada	Cond.	α
C1 8000-REIQB Churn	SHAP	Antigüedad en Meses es 1,0 (+2,973) Y Cargo Mensual es 69,95 (−0,221) Y Número de Referidos es 0,0 (+0,177) Y Cargos Totales de Larga Distancia es 22,5 (+0,146) Y Número de Dependientes es 0,0 (+0,116) Y ... (10 variables): Churn	10	—
	SHAP-L	Antigüedad en Meses es 1,0 (+2,973) Y Cargo Mensual es 69,95 (−0,221): Churn	2	—
	SHAP-R	SI Antigüedad en Meses es bajo Y Cargo Mensual es promedio ENTONCES Churn	2	0,906
C2 8849-AYPTR Churn	SHAP	Número de Referidos es 1,0 (+0,766) Y Antigüedad en Meses es 33,0 (−0,451) Y Cargo Mensual es 91,25 (+0,380) Y Número de Dependientes es 0,0 (+0,282) Y Promedio Mensual de GB Descargados es 17,0 (...) Y ... (10 variables): Churn	10	—
	SHAP-L	Número de Referidos es 1,0 (+0,766) Y Antigüedad en Meses es 33,0 (−0,451) Y Cargo Mensual es 91,25 (+0,380) Y Número de Dependientes es 0,0 (+0,282): Churn	4	—
	SHAP-R	SI Número de Referidos es promedio Y Antigüedad en Meses es promedio Y Cargo Mensual es por encima del promedio Y Número de Dependientes es promedio ENTONCES Churn	4	0,753
C3 2074-GKOWZ Churn	SHAP	Antigüedad en Meses es 2,0 (+2,264) Y Número de Referidos es 1,0 (+0,370) Y Cargo Mensual es 89,55 (+0,157) Y Número de Dependientes es 0,0 (+0,154) Y Cargos Totales de Larga Distancia es 17,72 (...) Y ... (10 variables): Churn	10	—
	SHAP-L	Antigüedad en Meses es 2,0 (+2,264) Y Número de Referidos es 1,0 (+0,370): Churn	2	—
	SHAP-R	SI Antigüedad en Meses es bajo Y Número de Referidos es promedio ENTONCES Churn	2	0,682
C4 5924-IFQTT No churn	SHAP	Número de Dependientes es 2,0 (−0,931) Y Número de Referidos es 4,0 (−0,772) Y Antigüedad en Meses es 13,0 (+0,175) Y Cargo Mensual es 72,8 (−0,156) Y Promedio Mensual de GB Descargados es 23,0 (...) Y ... (10 variables): No churn	10	—
	SHAP-L	Número de Dependientes es 2,0 (−0,931) Y Número de Referidos es 4,0 (−0,772) Y Antigüedad en Meses es 13,0 (+0,175): No churn	3	—
	SHAP-R	SI Número de Dependientes es promedio Y Número de Referidos es por encima del promedio Y Antigüedad en Meses es por debajo del promedio ENTONCES No churn	3	0,877

Para la primera regla de SHAP-R se muestra el contexto de la siguiente regla: **SI** Antigüedad en Meses es bajo **Y** Cargo Mensual es promedio **ENTONCES** Churn. En las Figuras 9 y 10 se observa el contexto estadístico.

(A) Panel resumen — Antigüedad en Meses

Indicador	Valor
Variable	Antigüedad en Meses
Valor	1.00
Etiqueta	bajo
μ	1.000
SHAP	+2.9727
Ranking SHAP	#1
Media	34.63
Mediana	33.00
Percentil	6%
Desviación estándar	24.07

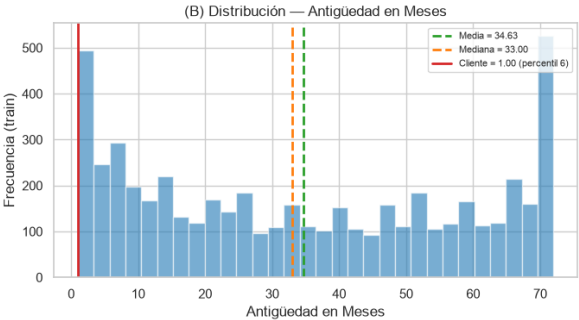


Figura 9. Contexto enriquecido del antecedente «Antigüedad en Meses es bajo»: función de pertenencia e histograma de entrenamiento. Fuente: elaboración propia.

(A) Panel resumen — Cargo Mensual

Indicador	Valor
Variable	Cargo Mensual
Valor	69.95
Etiqueta	promedio
μ	0.900
SHAP	-0.2209
Ranking SHAP	#2
Media	66.43
Mediana	73.15
Percentil	46%
Desviación estándar	29.76

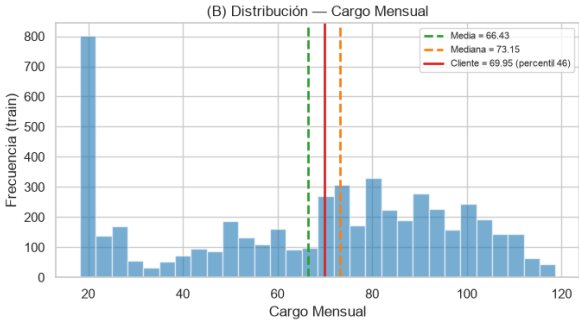


Figura 10. Contexto enriquecido del antecedente «Cargo Mensual es promedio»: función de pertenencia e histograma de entrenamiento. Fuente: elaboración propia.

2) Estudio 2: asignación de segmentos

En la Tabla 5 se muestra la comparación entre el resultado numérico de SHAP y SHAP-L y el resultado explicativo con lenguaje difuso de SHAP-R.

Tabla 5. Random Forest: comparativa de los formatos explicativos SHAP, SHAP-L y SHAP-R en la asignación de segmentos. Fuente: elaboración propia.

Instancia	Formato	Explicación generada	Cond.	α
S1 train #0 Cluster 1	SHAP	Total de Cargos Extra por Datos es 0,00 (+0,1556) Y Cargo Mensual es 60,00 (+0,1435) Y Descarga Promedio Mensual (GB) es 13,00 (+0,1162) Y Número de Referidos es 0,00 (+0,1082) Y Edad es 64,00 (+0,0285) Y Cargo Prom. Mensual por Larga Distancia es 21,64 (+0,0201) Y Número de Dependientes es 0,00 (+0,0131) Y Total de Cargos por Larga Distancia es 930,52 (-0,0090) Y Antigüedad en Meses es 43,00 (+0,0081) Y Total de Reembolsos es 0,00 (+0,0004): Cluster = 1	10	—
	SHAP-L	Total de Cargos Extra por Datos es 0,00 (+0,1556) Y Cargo Mensual es 60,00 (+0,1435) Y Descarga Promedio Mensual (GB) es 13,00 (+0,1162) Y Número de Referidos es 0,00 (+0,1082): Cluster = 1	4	—
	SHAP-R	SI Total de Cargos Extra por Datos es promedio Y Cargo Mensual es promedio Y Descarga Promedio Mensual (GB) es promedio Y Número de Referidos es promedio ENTONCES Cluster = 1	4	0,497

C. Comparación cuantitativa de los formatos explicativos

1) Interpretación de los resultados del Estudio 1

En ambos modelos, SHAP-R reduce la complejidad de las diez condiciones del formato completo a alrededor de tres, entre un 64 % y un 69 % menos de elementos que procesar, y a la vez eleva la consistencia por encima de la que alcanzan SHAP y SHAP-L (Figura 11).

Las Figuras 12 y 13 muestran la distribución de la fuerza de activación sobre las instancias de prueba. En ambos modelos la gran mayoría de los casos supera el

umbral $\alpha \geq 0,33$, lo que confirma que las reglas describen adecuadamente a los clientes que explican. La comparación entre ambas distribuciones favorece a LightGBM: su masa se concentra entre 0,45 y 0,75, con moda en torno a 0,60, mientras que la de Random Forest se desplaza hacia la izquierda, con moda cercana a 0,45 y una cola que alcanza valores de 0,20. Esta diferencia es coherente con las coberturas reportadas —0,978 frente a 0,895— y se explica por la mayor longitud media de las reglas del bosque aleatorio, ya que cada antecedente adicional multiplica un factor menor que la unidad en el cálculo de la activación.

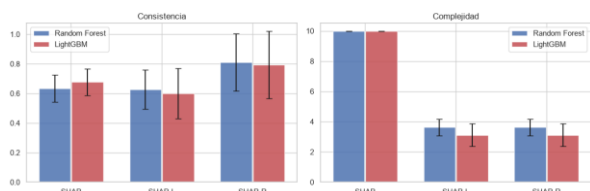


Figura 11. Consistencia y complejidad de los formatos SHAP, SHAP-L y SHAP-R. Fuente: elaboración propia.

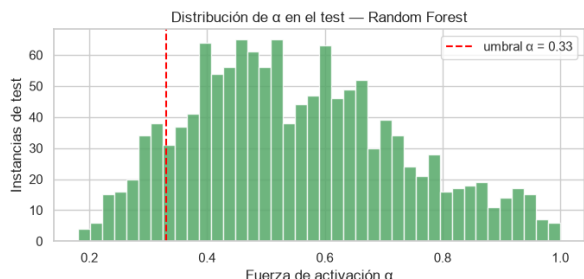


Figura 12. Distribución de la fuerza de activación y cobertura del modelo Random Forest. Fuente: elaboración propia.

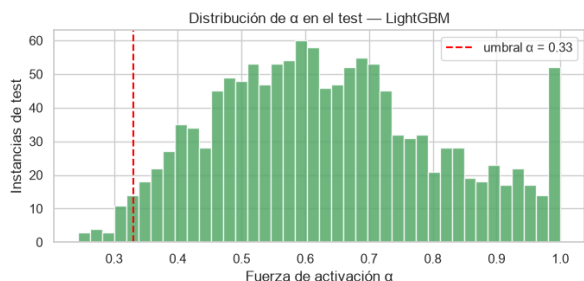


Figura 13. Distribución de la fuerza de activación y cobertura del modelo LightGBM. Fuente: elaboración propia.

2) Interpretación de los resultados del Estudio 2

La distribución agregada (Figura 14) es claramente bimodal: una masa principal entre 0,35 y 0,75 y un segundo pico pronunciado por encima de 0,90. El 92,8 % de las instancias supera el umbral $\alpha \geq 0,33$.

La descomposición por segmento (Figura 15) revela que esa bimodalidad no es una propiedad del método sino la superposición de dos comportamientos. Los segmentos 0 y 1 forman la masa central, con activaciones medias de 0,497 y 0,551 y coberturas del 87,2 % y 92,8 %; los segmentos 2 y 3 constituyen el pico derecho, con medias de 0,907 y 0,970 y coberturas del 100 % y 99,7 %. El orden reproduce el de la silueta individual, lo que confirma que las reglas describen mejor a los clientes de los segmentos mejor separados.

Ambos picos derechos deben interpretarse de forma distinta. El del segmento 2 es legítimo: los clientes sin internet registran exactamente cero gigabytes, valor que coincide con el vértice de su término. El del segmento 3, concentrado en $\alpha = 1,0$, procede de reglas de un solo antecedente asociado a una variable con partición colapsada, que se satisface para cualquier cliente. La activación máxima señala aquí ausencia de exigencia y no calidad de ajuste, lo que desaconseja emplear α como indicador único sin verificar el poder discriminante de los antecedentes.

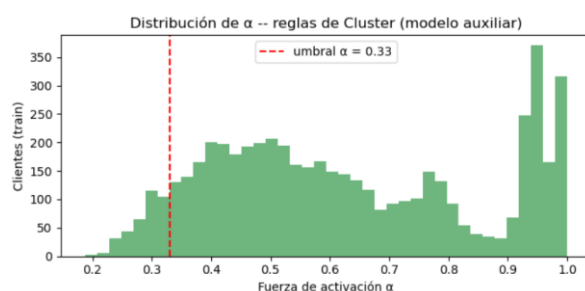


Figura 14. Distribución agregada de la fuerza de activación en la asignación de segmentos (Random Forest). Fuente: elaboración propia.

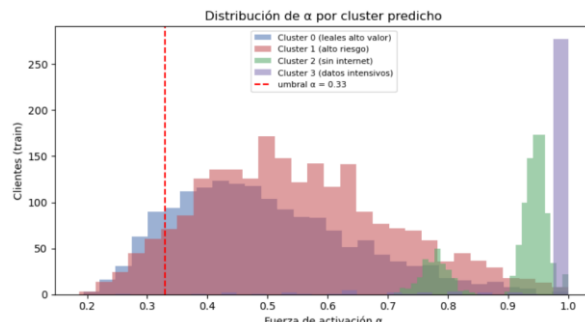


Figura 15. Distribución de la fuerza de activación por segmento (Random Forest). Fuente: elaboración propia.

VI. DISCUSIÓN

A. El valor añadido de la verbalización

El principal aporte de SHAP-R no radica en reducir el número de condiciones, sino en expresarlas mediante términos lingüísticos. Aunque SHAP-R y SHAP-L generan reglas de complejidad similar, SHAP-R alcanza una mayor consistencia textual. Esto indica que la mejora proviene de la verbalización difusa y no de una simplificación de las explicaciones.

B. Robustez de las funciones de pertenencia difusa

Las funciones de pertenencia difusas introducidas por SHAP-R hacen que pequeñas variaciones en los valores numéricos no impliquen necesariamente un cambio en la etiqueta lingüística asignada. Mientras una observación permanezca dentro del mismo conjunto difuso, la explicación se mantiene estable.

C. Eficiencia computacional

La incorporación de la capa lingüística no incrementa de forma apreciable el tiempo de generación de las explicaciones. Las diferencias observadas entre modelos responden al costo del cálculo de SHAP y no al proceso de verbalización. Por tanto, SHAP-R mantiene una latencia compatible con aplicaciones interactivas.

D. Limitación de las particiones difusas

La principal limitación identificada corresponde a variables con alta concentración de valores cero. En estos casos, las particiones basadas en cuartiles colapsan y generan términos lingüísticos poco informativos, produciendo antecedentes que se satisfacen para prácticamente todas las observaciones. Este resultado sugiere que la incorporación de un término específico para el valor cero resulta necesaria para aplicar el método en

dominios con inflación de ceros, como telecomunicaciones, banca o seguros.

E. Aporte de la capa de enriquecimiento

La regla difusa mejora la legibilidad, mientras que la capa de enriquecimiento recupera el nivel de detalle perdido al incorporar información estadística y la contribución relativa de cada antecedente. De esta forma, el usuario obtiene explicaciones comprensibles sin renunciar a la información necesaria para interpretar la importancia de cada variable. Al ser una capa exclusivamente descriptiva, no modifica las atribuciones originales ni introduce costo computacional adicional.

F. Limitaciones

El estudio presenta cuatro limitaciones principales.

Primero, el conjunto de datos es sintético y representa una empresa ficticia, por lo que los resultados no deben extrapolarse directamente a un entorno operativo.

Segundo, únicamente se consideran variables numéricas, excluyendo atributos categóricos potencialmente relevantes.

Tercero, el método utiliza una estrategia fija para generar las funciones de pertenencia, sin adaptarlas a la distribución particular de cada variable. En distribuciones complejas esta decisión puede reducir la capacidad discriminativa de algunos términos lingüísticos.

Cuarto, la segmentación adoptada no maximiza el coeficiente de silueta, y dos de los cuatro segmentos presentan una separación débil, lo que condiciona la interpretación de las reglas asociadas a ellos.

VII. RECOMENDACIONES

A. Para el despliegue en un sistema de apoyo a la decisión

1. **Priorizar LightGBM.** Con un desempeño equivalente al del bosque aleatorio, genera explicaciones un orden de magnitud más rápido y produce reglas más cortas, además de una cobertura superior. La diferencia no proviene de la capa explicativa sino de la profundidad de los árboles, por lo que limitar la profundidad del ensamble es una vía directa para reducir la latencia de explicación.
2. **Exponer siempre el bloque de atribución junto a la regla.** El porcentaje de contribución es lo que convierte la explicación en una guía de acción: permite distinguir un caso gobernado por un único factor de otro en el que la decisión se reparte entre varios, distinción que la conjunción lingüística no transmite.
3. **Usar el bloque de contexto difuso como indicador de confianza.** La distancia al término contiguo señala si la etiqueta asignada es estable o si el cliente se encuentra en una frontera. Los casos fronterizos son precisamente aquellos en los que la explicación puede cambiar ante variaciones mínimas y donde conviene un criterio humano adicional.
4. **Reportar la activación α al usuario final.** Las reglas por debajo del umbral de cobertura deben identificarse como tales en la interfaz, de modo que el analista sepa cuándo la explicación lingüística no

describe adecuadamente el caso y conviene recurrir al detalle numérico.

B. Para la aplicación del método a otros dominios

5. **Auditar la concentración modal antes de generar los términos.** Una proporción de moda superior a 0,75 anticipa el colapso de la partición. Este diagnóstico es inmediato y debe formar parte del análisis exploratorio, no descubrirse al inspeccionar las reglas.
6. **Incorporar un término explícito de tipo «es cero» o «ausente».** Es la vía natural para las variables con inflación de ceros, frecuentes en datos de comportamiento de clientes, financieros y de seguros. Alternativamente, excluir del conjunto de antecedentes candidatos aquellas variables cuya partición haya colapsado, evitando condiciones formalmente válidas y materialmente vacías.
7. **Eliminar la multicolinealidad antes de calcular las atribuciones.** Cuando dos variables son casi redundantes, el valor de Shapley reparte la contribución entre ambas y puede expulsar del conjunto seleccionado a una variable relevante por fragmentación de su importancia. El filtrado previo es aquí un requisito de la capa explicativa y no solo una práctica de modelado.
8. **Ajustar el consecuente al umbral operativo del modelo.** Si el punto de corte se ha desplazado respecto del valor por defecto, la regla debe reflejar la predicción efectiva bajo ese umbral; de lo contrario, la explicación puede contradecir la decisión que el sistema comunica.
9. **En problemas multiclase, tomar las atribuciones de la clase predicha.** De lo contrario, los antecedentes explican la propensión a una clase distinta de la asignada.

C. Para la investigación futura

Tres líneas se desprenden con naturalidad de las limitaciones señaladas. La primera es un estudio con usuarios que mida comprensión, tiempo de decisión y confianza, sustituyendo las métricas indirectas empleadas aquí por evidencia directa sobre el destinatario de la explicación. La segunda es la extensión del vocabulario lingüístico a variables categóricas y a distribuciones con inflación de ceros, que ampliaría el conjunto de dominios donde el método resulta aplicable. La tercera es un análisis de sensibilidad sobre los parámetros K , T y el umbral de activación, que en este trabajo se han fijado en los valores de la formulación original para permitir la comparación, pero que gobiernan directamente el compromiso entre cobertura, fidelidad y concisión.

VIII. CONCLUSIONES

Este trabajo propuso la **Explicación Difusa Enriquecida**, una extensión de SHAP-R que complementa las reglas lingüísticas con información estadística y de importancia local sin modificar las atribuciones originales ni incrementar el costo computacional. La propuesta fue evaluada en dos tareas de analítica de clientes en telecomunicaciones: la

predicción de fuga y la explicación de segmentos obtenidos mediante aprendizaje no supervisado.

Los resultados muestran que la mejora en interpretabilidad proviene de la **verbalización de las explicaciones y no únicamente de la reducción de su complejidad**. Aunque SHAP-R y la versión numérica truncada generan reglas de longitud similar, la representación mediante términos lingüísticos produce explicaciones más estables y consistentes, manteniendo una elevada fidelidad y cobertura respecto al modelo original. Estos resultados respaldan el uso de reglas difusas como una alternativa más comprensible que las explicaciones basadas exclusivamente en valores numéricos de SHAP.

La capa de enriquecimiento compensa la pérdida de información cuantitativa inherente a la verbalización al incorporar el percentil de la observación, la importancia relativa y el signo de cada antecedente, así como su proximidad a los límites de los conjuntos difusos. Al tratarse de una extensión estrictamente aditiva, preserva la explicación original y proporciona información adicional útil para apoyar la toma de decisiones.

La evaluación sobre el problema de segmentación confirma la **generalidad del método**. Las reglas generadas permiten describir los segmentos descubiertos mediante K-Means utilizando las variables originales del negocio, con niveles de fidelidad y consistencia comparables a los obtenidos en la clasificación binaria. Esto demuestra que la propuesta puede aplicarse tanto a problemas supervisados como a escenarios de aprendizaje no supervisado apoyados en clasificadores auxiliares.

Como principal aporte metodológico, este trabajo identifica y cuantifica una limitación no documentada previamente en SHAP-R. Las variables con alta concentración de valores idénticos generan particiones lingüísticas degeneradas que reducen el poder discriminativo de las reglas y pueden producir valores elevados de fuerza de activación sin reflejar una explicación realmente informativa. Este resultado pone de manifiesto la necesidad de adaptar la generación automática de conjuntos difusos para dominios con inflación de ceros, incorporando términos lingüísticos específicos para estos casos y revisando la métrica de activación para considerar el poder discriminativo efectivo de los antecedentes.

CONFLICTOS DE INTERÉS DE LOS AUTORES

Los autores declaran no tener conflicto de interés. No se recibió financiación externa para la realización de este trabajo.

CONTRIBUCIÓN DE LOS AUTORES

Extensión metodológica. Se propone la Explicación Difusa Enriquecida, una capa de extensión que, sin recalcular ni modificar el pipeline SHAP-Rule, complementa cada antecedente de la regla con cuatro bloques de contexto —básico, estadístico, de atribución y difuso— y dos visualizaciones normalizadas. Esta extensión responde a una ambigüedad concreta del formato lingüístico original: una etiqueta como «antigüedad es baja» no distingue si el cliente se sitúa

apenas por debajo del promedio o en el extremo de la distribución.

Transferencia de dominio. Se adapta y evalúa la metodología SHAP-Rule en el dominio de la predicción de fuga de clientes en telecomunicaciones, empleando un conjunto de datos público de 7 043 clientes y dos ensambles de árboles ampliamente utilizados en la literatura de churn.

Explicabilidad de segmentos no supervisados. Se extiende el método a la explicación de agrupamientos obtenidos mediante PCA y K-Means, generando reglas lingüísticas que enuncian por qué un cliente pertenece a un segmento. La aportación reside en trasladar la caracterización desde el espacio de componentes principales, opaco para el usuario, a las variables originales del negocio, y en verificar que la capa explicativa opera sin modificación estructural en un escenario multiclase.

Caracterización de un modo de fallo. Se documenta y caracteriza el colapso de la partición lingüística en variables con exceso de ceros, un modo de fallo que la formulación original anticipa de forma tentativa pero no cuantifica, y que en este conjunto de datos afecta a tres de las diez variables finales.

DISPONIBILIDAD DE DATOS Y CÓDIGO

El conjunto de datos empleado es de acceso público a través de [Kaggle: https://www.kaggle.com/datasets/shilongzhuang/telecom-customer-churn-by-maven-analytics](https://www.kaggle.com/datasets/shilongzhuang/telecom-customer-churn-by-maven-analytics)

LITERATURA CITADA

- Alexander, Z., Chau, D. H., & Saldaña, C. (2024). An interrogative survey of explainable AI in manufacturing. *IEEE Transactions on Industrial Informatics*, 20(5), 7069–7081. <https://doi.org/10.1109/TII.2024.3361489>
- Ali, N., Ahmed, L., & Gasmalla, T. A. A. (2026). *Data-driven telecom marketing optimization: A machine learning-based churn prediction and customer segmentation framework*. arXiv. <https://doi.org/10.48550/arXiv.2607.10260>
- Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion*, 99, 101805. <https://doi.org/10.1016/j.inffus.2023.101805>
- Antoni, T. G., Hakim, A. H., Hendartriany, R. N., Vyorra, A. V. E., & Septyanto, F. (2025). Precision-oriented churn prediction with a fine-tuned meta-learner stack model and SHAP: A case study on IBM Telco. *Indonesian Actuarial Journal*, 1(2), 137–151. <https://doi.org/10.65689/iajvol01no2pp137-151>
- Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges.

- WIREs Data Mining and Knowledge Discovery*, 13(2), e1484. <https://doi.org/10.1002/widm.1484>
- Cacao, J., Santos, J., & Antunes, M. (2025). Explainable AI for industrial fault diagnosis: A systematic review. *Journal of Industrial Information Integration*, 47, 100905. <https://doi.org/10.1016/j.jii.2025.100905>
- Cao, J., Zhou, T., & Zhi, S. (2024). Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis—A comprehensive review. *Information Sciences*, 662, 120212. <https://doi.org/10.1016/j.ins.2024.120212>
- Chen, H., Covert, I. C., Lundberg, S. M., & Lee, S.-I. (2023). Algorithms to estimate Shapley value feature attributions. *Nature Machine Intelligence*, 5, 590–601. <https://doi.org/10.1038/s42256-023-00657-x>
- Chen, X., Lv, B., Wang, M., Xiang, X., Wu, S., Luo, S., & Zhang, W. (2025). *A comprehensive analysis of churn prediction in telecommunications using machine learning* (arXiv:2509.22654). arXiv. <https://arxiv.org/abs/2509.22654>
- Cyril, H. P. (2026). An ensemble of advanced machine learning models for telecom customer churn forecasting. *Journal of Computer Science and Technology Studies*, 8(7), 103–111. <https://alkindipublisher.com/index.php/jcsts/article/view/12738>
- Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? In *Advances in Neural Information Processing Systems* (Vol. 35, pp. 507–520). Curran Associates. <https://arxiv.org/abs/2207.08815>
- Hooker, G., Mentch, L., & Zhou, S. (2021). Unrestricted permutation forces extrapolation: Variable importance requires at least one more model, or there is no free variable importance. *Statistics and Computing*, 31, 82. <https://doi.org/10.1007/s11222-021-10057-z>
- Khalyasmaa, A. I., Matrenin, P. V., & Eroshenko, S. A. (2025). Interpretable diagnostics with SHAP-Rule: Fuzzy linguistic explanations from SHAP values. *Mathematics*, 13(20), 3355. <https://www.mdpi.com/2227-7390/13/20/3355>
- Maven Analytics. (2022). *Telecom customer churn [Conjunto de datos]*. Kaggle. <https://www.kaggle.com/datasets/shilongzhuang/telecom-customer-churn-by-maven-analytics>
- Mendel, J. M., & Bonissone, P. P. (2021). Critical thinking about explainable AI (XAI) for rule-based fuzzy systems. *IEEE Transactions on Fuzzy Systems*, 29(12), 3579–3593. <https://doi.org/10.1109/TFUZZ.2021.3079503>
- Mettle, H. N.-A., Ayerko, E. H., & Appiahene, P. (2026). Machine learning-based customer churn prediction in telecommunication industry. *Applied Computational Intelligence and Soft Computing*, 2026, 8496186. <https://doi.org/10.1155/acis/8496186>
- Ouifak, H., & Idri, A. (2025). A comprehensive review of fuzzy logic based interpretability and explainability of machine learning techniques across domains. *Neurocomputing*, 647, 130602. <https://doi.org/10.1016/j.neucom.2025.130602>
- Sahakyan, M., Aung, Z., & Rahwan, T. (2021). Explainable artificial intelligence for tabular data: A survey. *IEEE Access*, 9, 135392–135422. <https://ieeexplore.ieee.org/document/9551946>
- Scornet, E., & Hooker, G. (2026). Theory of random forests. *Annual Review of Statistics and Its Application*, 13, 99–121. <https://doi.org/10.1146/annurev-statistics-112723-034707>
- Shaikhshurab, M. A., & Magadum, P. (2024). *Enhancing customer churn prediction in telecommunications: An adaptive ensemble learning approach* (arXiv:2408.16284). arXiv. <https://arxiv.org/abs/2408.16284>
- Sikri, A., Jameel, R., Idrees, S. M., & Kaur, H. (2024). Enhancing customer retention in telecom industry with machine learning driven churn prediction. *Scientific Reports*, 14, 13097. <https://doi.org/10.1038/s41598-024-63750-0>
- So, B. (2024). Enhanced gradient boosting for zero-inflated insurance claims and comparative analysis of CatBoost, XGBoost, and LightGBM. *Scandinavian Actuarial Journal*, 2024(10), 1013–1035. <https://doi.org/10.1080/03461238.2024.2365390>
- Verdinelli, I., & Wasserman, L. (2024). Feature importance: A closer look at Shapley values and LOCO. *Statistical Science*, 39(4), 623–636. <https://doi.org/10.1214/24-STS937>
- Vishakha, F. (2026). *A comparative analysis of machine learning models for customer churn prediction in the telecommunications industry: A business analytics perspective*. College of Business and Economics, California State University, Fullerton. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6819479
- Zerine, I., Islam, M. M., Khan, M. A. U., Chy, M. A. R., Saimon, A. S. M., Manik, M. M. T. G., & Wata, C. (2026). Explainable churn prediction in telecom with tabular ML: Five model benchmark and SHAP analysis. *Discover Artificial Intelligence*, 6, 263. <https://doi.org/10.1007/s44163-026-00983-0>