

RIESGOS FINANCIEROS

Notas Técnicas:

Análisis de Sentimientos en Depósitos a la Vista

Febrero 05, 2026

RIESGOS FINANCIEROS NOTAS TÉCNICAS

Análisis de Sentimientos en Depósitos a la Vista

PÁG. 2

Antecedentes

PÁG. 3

¿Qué son los Modelos de Sentimientos?

PÁG. 4

¿Cómo se calculan los Modelos de Análisis de Sentimientos?

PÁG 5

Modelos de Machine Learning para clasificación

PÁG 6

Metodología

PÁG 14

Aplicaciones y Usos

PÁG 15

Supuestos Técnicos

PÁG 16

Riesgos y Limitantes

PÁG 17

Referencias



OBJETIVO GENERAL

El objetivo de esta nota técnica es analizar la relación entre las opiniones sobre la banca en redes sociales y la sensibilidad de sus depósitos. Se busca desarrollar un marco que permita anticipar salidas aceleradas de depósitos en respuesta a eventos que generen percepciones de fragilidad o tensión en el sector. Este enfoque facilitará la implementación de medidas preventivas en la gestión del balance y subraya la importancia de considerar las dinámicas sociales de las plataformas digitales como una nueva dimensión en la evaluación del riesgo de liquidez.

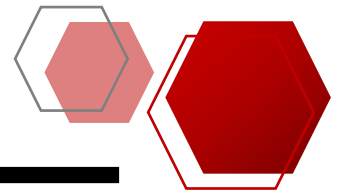
INTRODUCCIÓN

En los últimos años, las redes sociales han adquirido un papel importante en la percepción pública sobre la salud financiera de los bancos. Eventos como las corridas bancarias de Silicon Valley Bank (SVB), Signature Bank y First Republic Bank durante el 2023 han evidenciado cómo las plataformas digitales pueden amplificar el riesgo sistémico mediante la rápida propagación de información (y desinformación), modificando el comportamiento de los depositantes en tiempos de tensión financiera.

Diversos estudios recientes (Cookson et al., 2023; Gam et al., 2024; ECB, 2023) evidencian que la exposición de los bancos a las redes sociales, en particular a X (antes Twitter), pueden acelerar los retiros de depósitos cuando existen señales de deterioro, incluso antes de que los indicadores financieros tradicionales justifiquen dichos movimientos. Esto ocurre porque los depositantes no solo reaccionan a la información directa que reciben del banco, sino también a sus expectativas sobre el comportamiento de otros depositantes. En este contexto, las redes sociales permiten observar en tiempo real el sentimiento general de los usuarios, amplificando dinámicas de contagio y coordinación.

ANTECEDENTES

En años recientes, la expansión del acceso a internet y la aparición de nuevos canales digitales, en particular los medios informativos en línea y las redes sociales, han generado un crecimiento acelerado en la producción de datos, especialmente de tipo no estructurado. Este entorno ha impulsado el desarrollo de tecnologías y metodologías avanzadas capaces de procesar grandes volúmenes de información en formatos diversos, permitiendo extraer señales relevantes con mayor oportunidad.



El Banco de México ha documentado este cambio estructural en su Reporte de Estabilidad Financiera (diciembre 2024), destacando que el análisis de texto proveniente de noticias y redes sociales ofrece una ventana privilegiada para monitorear expectativas, percepciones de riesgo y narrativas que pueden incidir en la estabilidad financiera. En particular, el uso de técnicas de Procesamiento de Lenguaje Natural (NLP) ha permitido construir indicadores de alta frecuencia como el Índice de Incertidumbre de Política Económica (EPU) y el Índice de Riesgo Bancario, ambos diseñados para detectar señales tempranas de tensiones económicas o financieras a partir de la dinámica informativa digital.

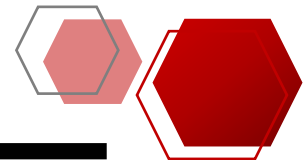
Estos desarrollos se apoyan en una literatura creciente. Por un lado, Baker et al. (2016) demostraron el valor de cuantificar la incertidumbre económica mediante el análisis sistemático de noticias; por otro, estudios más recientes, como Cookson et al. (2023), han evidenciado el papel crítico de las redes sociales en la aceleración de episodios de estrés bancario. A nivel internacional, iniciativas como el Proyecto Elipse del BIS han reforzado el uso de datos no estructurados dentro de sistemas de alerta temprana aplicados al sector financiero, subrayando su utilidad para identificar vulnerabilidades emergentes.



En conjunto, estos antecedentes muestran que la integración de datos digitales y metodologías de NLP constituyen un componente esencial para la supervisión moderna de riesgos. Además, establecen la base técnica y conceptual para desarrollar modelos de aprendizaje automático capaces de anticipar episodios de estrés en los depósitos, especialmente cuando estos pueden originarse o amplificarse en el entorno informativo de redes sociales y medios digitales.

¿QUÉ SON LOS MODELOS DE SENTIMIENTOS?

Uno de los modelos más usados dentro del campo de Procesamiento de Lenguaje Natural (NLP, por sus siglas en inglés) son los modelos de análisis de sentimiento, los cuales son herramientas computacionales diseñadas para identificar, clasificar y cuantificar las opiniones, emociones o actitudes expresadas en textos escritos en lenguaje natural. Su finalidad principal es determinar la polaridad del contenido, típicamente clasificada como positiva, negativa o neutral, aunque enfoques más avanzados permiten medir la intensidad del sentimiento o identificar



Estos modelos utilizan técnicas estadísticas, de aprendizaje automático y de aprendizaje profundo para extraer información estructurada a partir de grandes volúmenes de texto no estructurado, como publicaciones en redes sociales, foros, reseñas o noticias (Jurafsky & Martin, 2023).

Debido a esta capacidad para capturar percepciones de manera oportuna, los modelos de sentimiento se han vuelto fundamentales en aplicaciones donde las actitudes de los usuarios pueden tener impactos inmediatos, como los mercados financieros, la reputación institucional, el análisis de riesgo o la detección temprana de episodios de estrés. En estos contextos, el análisis de sentimientos ofrece una aproximación sistemática para estudiar cómo evoluciona la percepción pública y, sobre todo, cómo esta puede traducirse en comportamientos que afecten la estabilidad o el funcionamiento de un sistema determinado.



¿CÓMO SE CALCULAN LOS MODELOS DE ANÁLISIS DE SENTIMIENTOS?

El cálculo de los modelos de análisis de sentimiento consiste en una serie de etapas metodológicas que permiten transformar texto no estructurado en métricas cuantitativas que representan la polaridad o intensidad del sentimiento. Aunque los procedimientos específicos dependen del enfoque utilizado, la literatura coincide en que el proceso general incluye: (i) preprocesamiento del texto, (ii) representación numérica del lenguaje y (iii) estimación o inferencia del sentimiento mediante un modelo formal (Liu, 2012; Jurafsky & Martin, 2023).

Preprocesamiento del texto

El preprocesamiento es una etapa fundamental cuyo objetivo es limpiar y normalizar los textos para reducir ruido y mejorar el desempeño del modelo. Entre las técnicas más comunes se encuentran la eliminación de signos de puntuación, URLs o links, caracteres especiales, la conversión a minúsculas, la eliminación de palabras vacías (stopwords), y la lematización o stemming (normalización) de las palabras.

En el caso de textos provenientes de redes sociales, el preprocesamiento adquiere especial relevancia debido al uso frecuente de abreviaturas, emoticonos, hashtags y lenguaje informal, los cuales pueden contener información relevante sobre el sentimiento y, por tanto, deben ser tratados cuidadosamente (Taboada et al., 2011).

Cálculo del sentimiento en modelos basados en léxicos

En los modelos basados en léxicos, el cálculo del sentimiento se realiza asignando a cada palabra del texto un valor numérico previamente definido en un diccionario de sentimiento. Estos valores suelen representar polaridad (positiva o negativa) o intensidad emocional. El sentimiento total de un texto se obtiene agregando los valores de las palabras que lo componen, generalmente mediante sumas o promedios (Liu, 2012).

Formalmente, el puntaje de sentimiento de un texto S puede expresarse como:

$$S = \sum_{i=1}^N w_i$$

Donde w_i representa la puntuación de sentimiento asociada a la palabra i y N es el número total de palabras con carga emocional en el texto. En aplicaciones financieras, es común el uso de diccionarios especializados que ajustan la semántica al contexto económico, dado que palabras con connotación negativa en lenguaje general pueden no serlo en documentos financieros (Loughran & McDonald, 2011).

Cálculo del sentimiento en modelos de aprendizaje automático supervisado

En los modelos supervisados, el cálculo del sentimiento se basa en el entrenamiento de un algoritmo de clasificación a partir de un conjunto de textos previamente etiquetados. En una primera etapa, los textos se transforman en vectores numéricos mediante técnicas de representación como bolsa de palabras (bag-of-words), TF-IDF o n-grams (Pang et al., 2002).

Posteriormente, el modelo estima una función que asigna una probabilidad a cada clase de sentimiento. En el caso de un modelo de clasificación binaria, el sentimiento de un texto se determina como:

$$P(y = 1 | x) = f(x; \theta)$$

Donde x representa el vector de características del texto, θ los parámetros estimados del modelo y y la categoría de sentimiento. El resultado puede expresarse como una etiqueta discreta (positivo, negativo o neutral) o como una probabilidad continua, lo cual resulta útil para la construcción de indicadores agregados de sentimiento.

Cálculo del sentimiento en modelos de aprendizaje profundo

Los modelos de aprendizaje profundo emplean representaciones distribuidas del lenguaje, conocidas como embeddings, que capturan el significado semántico y el contexto de las palabras. En particular, los modelos basados en arquitecturas transformer procesan el texto completo de manera bidireccional, permitiendo que el significado de cada palabra dependa del contexto en el que aparece (Devlin et al., 2019).

En estos modelos, el cálculo del sentimiento se obtiene a partir de la salida de la red neuronal, usualmente una capa de clasificación que produce una distribución de probabilidad sobre las categorías de sentimiento. El puntaje final puede interpretarse como una probabilidad, un valor continuo o una clasificación discreta, dependiendo de la aplicación específica.

Agregación y construcción de indicadores de sentimiento

Una vez calculado el sentimiento a nivel de comentario individual, los resultados pueden agregarse para construir indicadores temporales o institucionales. Por ejemplo, el sentimiento promedio diario, el porcentaje de comentarios negativos o un índice ponderado por volumen de interacción. Este tipo de agregación permite vincular el sentimiento extraído de redes sociales con variables económicas y financieras observables, facilitando su uso en análisis estadísticos y econométricos (Tetlock, 2007).

En el contexto bancario, estos indicadores pueden interpretarse como medidas indirectas de percepción, confianza o riesgo reputacional, lo que justifica su análisis conjunto con variables como la liquidez de los depósitos.

MODELOS DE MACHINE LEARNING PARA CLASIFICACIÓN

Enfoque general de Machine Learning

Los modelos de Machine Learning se emplean para identificar relaciones complejas entre un conjunto de variables explicativas y una variable objetivo, especialmente cuando dichas relaciones no pueden representarse adecuadamente mediante supuestos paramétricos simples. Estos enfoques permiten extraer patrones a partir de los datos de manera flexible, lo que los hace particularmente útiles en aplicaciones económicas y financieras donde las interacciones entre variables suelen ser no lineales y de alta dimensionalidad.

Clasificación supervisada

Dentro del conjunto de técnicas de Machine Learning, la clasificación supervisada se centra en la asignación de observaciones a categorías discretas a partir de información histórica previamente etiquetada. El modelo aprende una función que relaciona las variables explicativas con una variable objetivo categórica, ajustando sus parámetros durante el proceso de entrenamiento para minimizar los errores de clasificación observados en los datos históricos. Este enfoque ha sido ampliamente documentado en la literatura de aprendizaje estadístico como una alternativa flexible a los modelos tradicionales cuando se trabaja con conjuntos de datos complejos (Hastie, Tibshirani y Friedman, 2009).

En el ámbito financiero, este tipo de modelos se utiliza para anticipar distintos comportamientos económicos, tales como cambios en la calidad crediticia, decisiones de los agentes o variaciones en variables de balance. En este estudio, el problema se plantea como uno de clasificación binaria, donde el objetivo es predecir la dirección del cambio en los saldos de los depósitos bancarios en un periodo determinado.

Indicadores de sentimiento como variables explicativas

Como se mencionó previamente, los datos provenientes de redes sociales son, por naturaleza, no estructurados. A través del análisis de sentimientos, esta información se transforma en medidas cuantitativas que resumen el tono general de los comentarios emitidos por los usuarios en un periodo específico. Los indicadores de sentimiento contruidos a partir de este proceso permiten capturar la percepción pública hacia una institución financiera de manera agregada y comparable en el tiempo.

Diversos estudios han mostrado que este tipo de indicadores contienen información relevante para explicar y anticipar el comportamiento de variables financieras, lo que respalda su uso como variables explicativas dentro de modelos predictivos (Tetlock, 2007). En este sentido, los indicadores de sentimiento funcionan como un puente entre la información cualitativa de redes sociales y los modelos cuantitativos utilizados en el análisis financiero.

Modelos de clasificación en aplicaciones financieras

La literatura en Machine Learning financiero ha explorado una amplia variedad de algoritmos de clasificación, entre los que se encuentran la regresión logística, los árboles de decisión, los métodos de ensamble, las redes neuronales y las máquinas de vectores de soporte. La selección del modelo más adecuado depende de las características específicas de los datos, incluyendo el tamaño de la muestra, el grado de correlación entre las variables explicativas y la posible presencia de relaciones no lineales con la variable objetivo.

Ventajas del enfoque de Machine Learning en este estudio

Los modelos de Machine Learning presentan ventajas relevantes frente a enfoques paramétricos tradicionales, ya que no requieren supuestos estrictos sobre la distribución de los datos y permiten capturar patrones complejos de manera flexible. Estas características resultan especialmente valiosas cuando se utilizan indicadores derivados de redes sociales, los cuales suelen presentar alta volatilidad y dependencia entre sí. En este contexto, los modelos de clasificación supervisada proporcionan un marco metodológico adecuado para integrar información de sentimiento y variables financieras, sirviendo como base conceptual para el modelo propuesto en la siguiente sección.

METODOLOGÍA

Enfoque General

Para la estimación del sentimiento contenido en los comentarios de redes sociales se emplea un modelo de aprendizaje profundo basado en arquitecturas de tipo transformer (modelo de redes neuronales que utiliza mecanismos de atención para capturar relaciones semánticas en textos cortos y largos), ajustado específicamente para la tarea de clasificación de sentimiento en texto corto. Este enfoque permite capturar relaciones semánticas complejas y dependencias contextuales que no pueden ser modeladas adecuadamente mediante enfoques tradicionales basados en conteo de palabras o clasificadores lineales.

El resultado del modelo se utiliza para construir indicadores cuantitativos de sentimiento que posteriormente se analizan en conjunto con los saldos de los depósitos bancarios.

Arquitectura del modelo de aprendizaje profundo.

El modelo utilizado se basa en una arquitectura transformer preentrenada, la cual representa cada texto de entrada como una secuencia de tokens (t_1, t_2, \dots, t_n) . Cada token es transformado inicialmente en una representación vectorial mediante una capa de embeddings (vectores numéricos que representan palabras o tokens capturando similitudes semánticas y relaciones contextuales) incorporando información semántica y posicional.

Formalmente, la entrada del modelo puede representarse como:

$$X = [x_1, x_2, \dots, x_n], \quad x_i \in \mathbb{R}^d$$

Donde d es la dimensión del espacio de embeddings. Estas representaciones son procesadas a través de múltiples capas de autoatención (self-attention), las cuales permiten que cada token incorpore información del contexto completo del texto.

En cada capa de atención, el mecanismo de scaled dot-product attention (el mecanismo pondera la importancia de cada token en función de su similitud con otros tokens dentro de una secuencia) se define como:

$$Attention(Q, K, V) = softmax \left(\frac{QK^T}{\sqrt{d_K}} \right) V$$

Donde Q, K y V representan las matrices de consultas (queries), claves (keys) y valores (values), respectivamente, y d_K es la dimensión de las claves. Este mecanismo permite ponderar dinámicamente la importancia de cada palabra en función de su contexto.

Capa de clasificación de sentimiento

Para la tarea de análisis de sentimiento, se utiliza la representación contextual agregada del texto, típicamente asociada a un token especial de clasificación. Esta representación final $h \in \mathbb{R}^d$ se introduce en una capa completamente conectada (feed-forward) que actúa como clasificador:

$$z = Wh + b$$

Donde $W \in \mathbb{R}^{K \times d}$ es la matriz de pesos, $b \in \mathbb{R}^K$ el vector de sesgos y K el número de clases de sentimiento consideradas (positivo, neutral y negativo).

La probabilidad asociada a cada clase de sentimiento se obtiene aplicando una función softmax (función de activación que convierte puntajes del modelo en probabilidades normalizadas sobre un conjunto de clases):

$$P(S = k | x) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}}$$

El sentimiento asignado a cada comentario corresponde a la clase con mayor probabilidad estimada, aunque las probabilidades continuas se pueden conservar para la construcción de indicadores agregados.

Función de pérdida y entrenamiento del modelo

Durante la etapa de ajuste del modelo (fine-tuning), los parámetros se estiman minimizando una función de pérdida de entropía cruzada categórica:

$$\mathcal{L} = - \sum_{k=1}^K y_k \log (P(S = k | x))$$

Donde y_k representa la etiqueta verdadera del sentimiento. El proceso de optimización se realiza mediante algoritmos de descenso de gradiente estocástico adaptativo, actualizando los pesos de la red para maximizar la capacidad predictiva del modelo sobre datos no observados.

Construcción del puntaje de sentimiento

Una vez entrenado el modelo, se obtiene para cada comentario i un vector de probabilidades (p_i^+, p_i^0, p_i^-) . A partir de estas probabilidades se define un puntaje continuo de sentimiento como:

$$S_i = p_i^+ - p_i^-$$

Este puntaje permite captar no solo la dirección del sentimiento, sino también su intensidad relativa, facilitando su agregación temporal.

Agregación temporal del sentimiento

Los puntajes individuales se agregan a nivel temporal para construir indicadores de sentimiento institucional. Para un periodo t , el indicador agregado se define como:

$$\bar{S}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} S_{t,i}$$

Donde N_t es el número de comentarios observados en el periodo t . Este indicador resume la percepción promedio expresada en redes sociales y se utiliza como variable explicativa en el análisis de su relación con la liquidez de los depósitos bancarios.

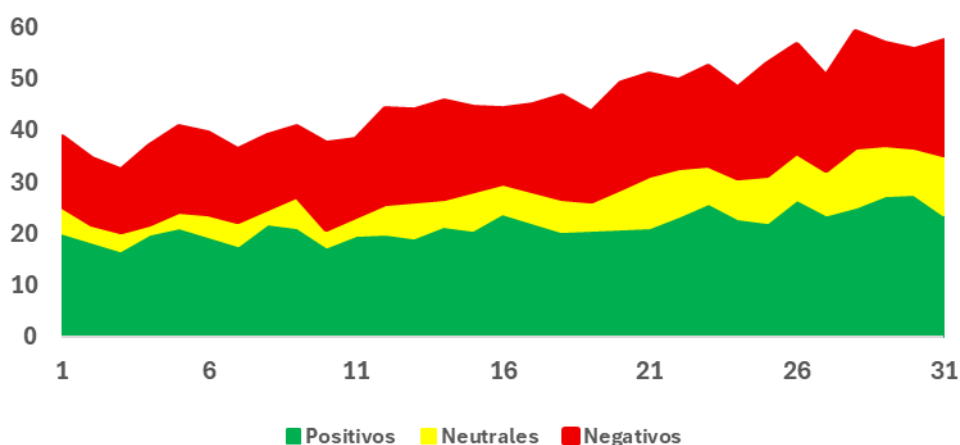
Construcción de los indicadores de sentimiento

El modelo de aprendizaje profundo descrito en la sección anterior permite asignar a cada comentario individual una clasificación discreta de sentimiento (positivo, neutral o negativo), así como un puntaje continuo asociado a la intensidad del sentimiento. Si bien estos resultados son informativos a nivel individual, para su análisis conjunto con variables financieras agregadas, como el comportamiento de los depósitos bancarios, es necesario transformarlos en indicadores cuantitativos que resuman la percepción expresada en redes sociales a lo largo del tiempo.

Sea t un periodo de agregación temporal (por ejemplo, diario o mensual) y sea N_t el número total de comentarios observados en dicho periodo. Para cada comentario i , el modelo produce una clasificación discreta de sentimiento y un puntaje continuo S_i .

A partir de la clasificación discreta generada por el modelo, se obtiene el conteo de comentarios clasificados como positivos, neutrales y negativos en cada periodo t . Estos conteos permiten describir de manera preliminar la distribución del sentimiento expresado en redes sociales y constituyen el insumo básico para la construcción de los indicadores agregados.

Figura 1. Conteo de comentarios por categoría de sentimiento



La figura 1 muestra la evolución temporal del número de comentarios clasificados como positivos, neutrales y negativos identificados en cada periodo por el modelo de análisis de sentimiento.

A partir de los resultados anteriores, se construyen indicadores agregados que permiten resumir distintas dimensiones del sentimiento expresado en redes sociales y facilitan su análisis cuantitativo.

Proporción de comentarios negativos

El primer indicador corresponde a la proporción de comentarios clasificados como negativos respecto al total de comentarios en el periodo t . Este indicador permite capturar episodios de mayor percepción adversa o descontento expresado en redes sociales.

Sentimiento promedio

El sentimiento promedio resume el tono general de los comentarios observados en el periodo t , utilizando el puntaje continuo estimado por el modelo de aprendizaje profundo. Se define como:

$$\bar{S}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} S_{t,i}$$

Donde \bar{S}_t representa el puntaje de sentimiento promedio del periodo t . Valores mayores de este indicador reflejan un tono más positivo, mientras que valores menores reflejan una percepción más negativa.

Índice neto de sentimiento

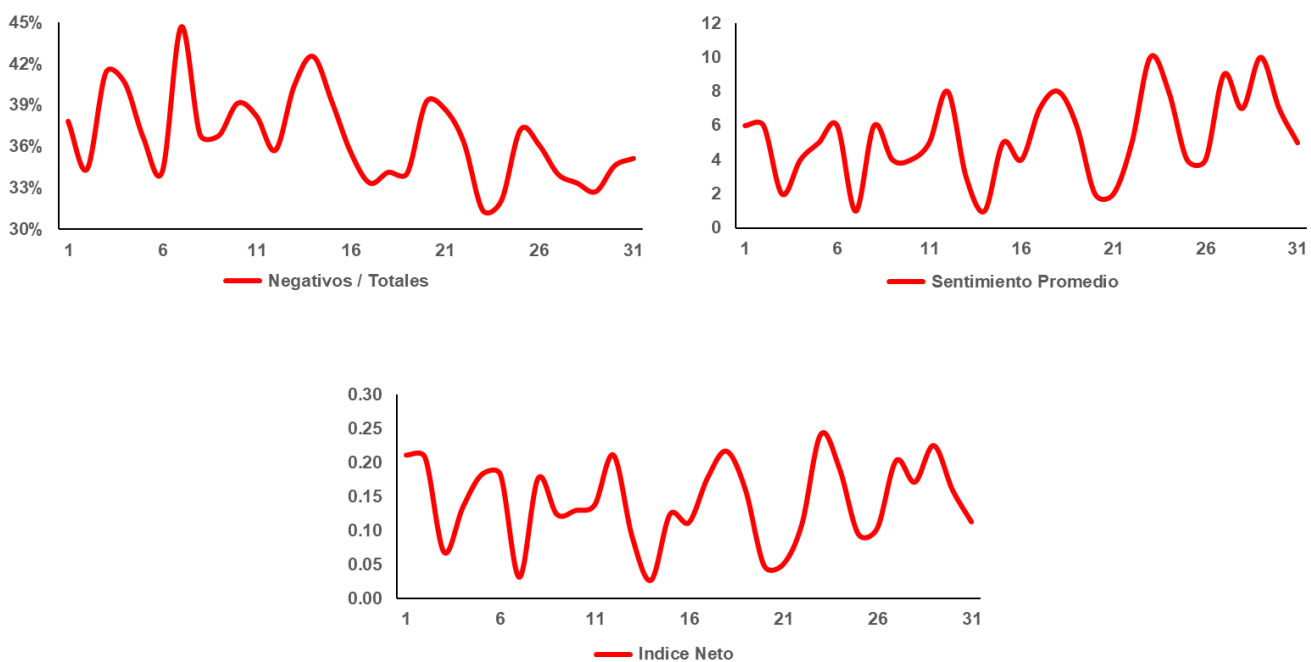
Este índice mide el balance entre opiniones positivas y negativas, proporcionando una medida sintética de la percepción expresada en redes sociales. Se define como:

$$Indice\ Neto_t = \frac{Positivos_t - Negativos_t}{Positivos_t + Negativos_t}$$

Este indicador toma valores en el intervalo de $[-1, 1]$, donde valores positivos indican predominancia de comentarios favorables y valores negativos reflejan predominancia de opiniones desfavorables.

A continuación se observan los indicadores descritos previamente:

Figura 2. Evolución temporal de los indicadores de sentimiento



La Figura 2 muestra la evolución de los indicadores construidos a partir de los comentarios en redes sociales: la proporción de comentarios negativos, el sentimiento promedio y el índice neto de sentimiento. Estas series permiten analizar la dinámica del sentimiento a lo largo del periodo de estudio e identificar cambios en el tono general de las opiniones expresadas por los usuarios.

Los indicadores definidos en esta sección se emplean como variables explicativas en el modelo de clasificación basado en Support Vector Machines, descrito en la sección siguiente. **Este modelo no procesa texto de manera directa; en su lugar, utiliza los indicadores agregados derivados del análisis de sentimiento.** Al integrarse con la misma frecuencia temporal que las variables financieras, dichos indicadores permiten evaluar si la información contenida en el sentimiento expresado en redes sociales contribuye a distinguir entre periodos de aumento o disminución en el saldo de los depósitos bancarios.

La figura 3 resume la metodología propuesta para el análisis de sentimiento:

Figura 3. Flujo general de la metodología de análisis de sentimiento.



Se observa el flujo general del modelo de análisis de sentimiento, desde la recolección de comentarios provenientes de redes sociales, foros y noticias, hasta la estimación de probabilidades de sentimiento mediante una arquitectura basada en Transformers. Dichas probabilidades se agregan temporalmente para construir indicadores de percepción institucional, los cuales se emplean como variables explicativas en el análisis de la sensibilidad de los depósitos bancarios.

Modelo de clasificación basado en Support Vector Machines

Con el objetivo de evaluar la relación entre los indicadores de sentimiento construidos en la sección anterior y los cambios en el saldo de los depósitos bancarios, se emplea un modelo de clasificación basado en Support Vector Machines (SVM). Este enfoque permite evaluar si la información contenida en los indicadores de redes sociales contribuye a discriminar entre periodos de aumento o disminución en el saldo de los depósitos.

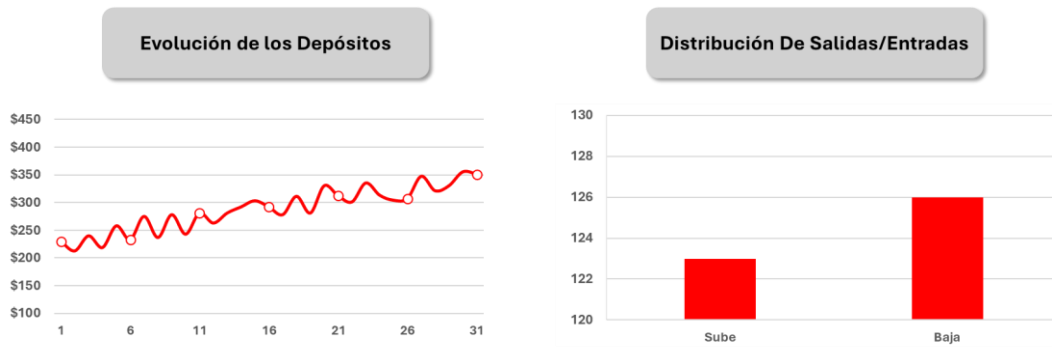
Definición de variables

Sea t un periodo de observación temporal. Se define la variable dependiente categórica como:

$$y_t = \begin{cases} 1, & \text{si el saldo de los depósitos aumenta en el periodo } t \\ 0, & \text{si el saldo de los depósitos disminuye o no aumenta} \end{cases}$$

Esta variable captura la dirección del cambio en el saldo de los depósitos y permite formular el problema como una tarea de clasificación binaria.

Con el fin de ilustrar el proceso de transformación de la serie continua del saldo de los depósitos en una variable categórica, se presenta a continuación una representación gráfica de dicha transformación.

Figura 4. Transformación del saldo de los depósitos en variable binaria


La figura 4 muestra la evolución temporal del saldo de los depósitos y la distribución de los periodos clasificados como aumentos o disminuciones. La clasificación se realiza a partir de la variación del saldo entre periodos consecutivos, asignando la categoría de aumento cuando la variación es positiva y de disminución cuando es negativa.

El vector de variables explicativas $x_t \in \mathbb{R}^p$ está compuesto por los indicadores de sentimiento (sentimiento promedio del periodo, proporción de comentarios negativos respecto al total, índice de sentimiento neto, etc.) descritos previamente. Los indicadores se representan formalmente como:

$$x_t = (s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(p)})$$

Donde p es la cantidad de indicadores construidos a partir del análisis de sentimiento.

Formulación matemática del modelo SVM

El modelo SVM busca encontrar un hiperplano que separe óptimamente las observaciones pertenecientes a las dos clases definidas por y_t . En su forma lineal, el clasificador se expresa como:

$$f(x) = w^T x + b$$

Donde w es el vector de pesos y b el término de sesgo. La regla de clasificación es:

$$\hat{y} = \begin{cases} 1, & \text{si } f(x) \geq 0 \\ 0, & \text{si } f(x) < 0 \end{cases}$$

El problema de optimización asociado al SVM con margen suave se define como:

$$\min_{w, b, \xi} = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

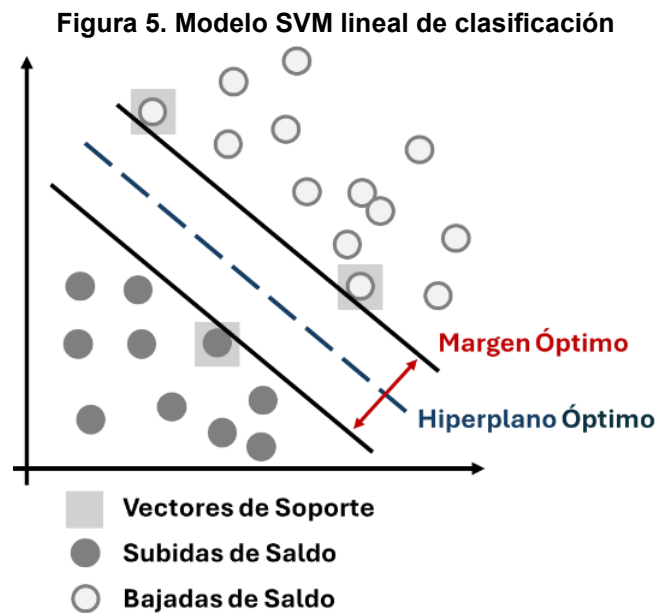
Sujeto a:

$$y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

Donde:

- ξ_i son variables de holgura que permiten errores de clasificación.
- $C > 0$ es el parámetro de regularización que controla el compromiso entre maximizar el margen y penalizar errores.

El modelo en su versión lineal se puede visualizar de la siguiente manera:



En la figura 5 podemos observar un modelo de SVM lineal de 2 dimensiones, los ejes representan un set de 2 variables que tienen una relación directa con la subida o baja de los saldos. La idea del modelo es encontrar un hiperplano que permita separar las observaciones correctamente. Los vectores de soporte, que vienen representados por los cuadros grises, sirven para definir cuál será el margen más largo entre las 2 clases.

Extensión no lineal mediante funciones kernel

Dado que la relación entre los indicadores de sentimiento y la variación de los depósitos puede ser no lineal, se considera la extensión del modelo mediante una función kernel (mide la similitud entre observaciones en un espacio transformado) $K(x_i, x_j)$, que permite proyectar los datos a un espacio de mayor dimensión sin calcular explícitamente dicha transformación.

Una función kernel comúnmente utilizada es el kernel de base radial (RBF):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

Donde γ controla la influencia relativa de cada observación.

Entrenamiento y validación del modelo

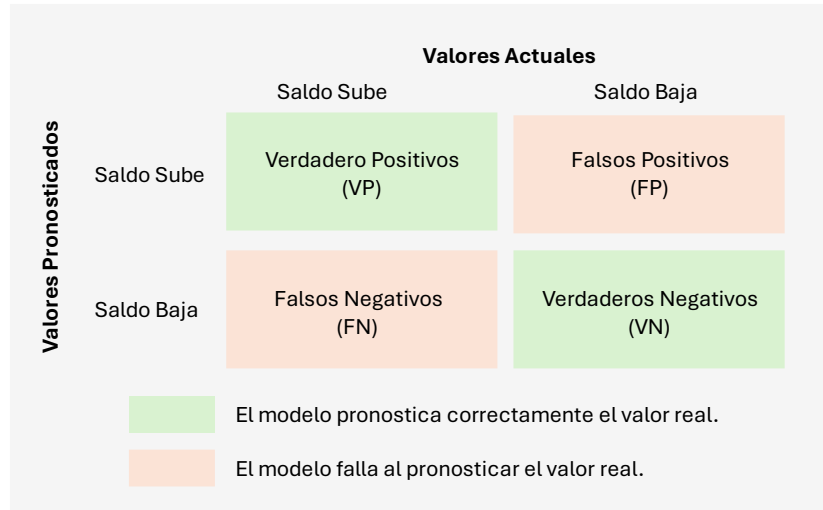
El conjunto de datos se divide en subconjuntos de entrenamiento y prueba, preservando la estructura temporal para evitar filtraciones de información. El modelo se entrena utilizando el conjunto de entrenamiento y se evalúa sobre el conjunto de prueba.

La selección de hiperparámetros (valores externos al modelo que regulan su desempeño y se seleccionan durante el proceso de validación) (C, γ) se realiza mediante validación cruzada, maximizando el desempeño predictivo del modelo.

Evaluación del desempeño del modelo (Matriz de confusión)

El desempeño del clasificador se evalúa a partir de la matriz de confusión, la cual resume las predicciones del modelo en comparación con los valores observados:

Figura 6. Matriz de confusión de modelo de clasificación



Esta matriz (figura 6) permite identificar errores de clasificación asimétricos, particularmente relevantes en el análisis de riesgo financiero, donde las consecuencias de una clasificación incorrecta pueden diferir según el tipo de error.

Métricas de evaluación

A partir de la matriz de confusión se calculan las siguientes métricas:

- Accuracy (exactitud)

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN}$$

- Precisión (Precision)

$$Precision = \frac{VP}{VP + FP}$$

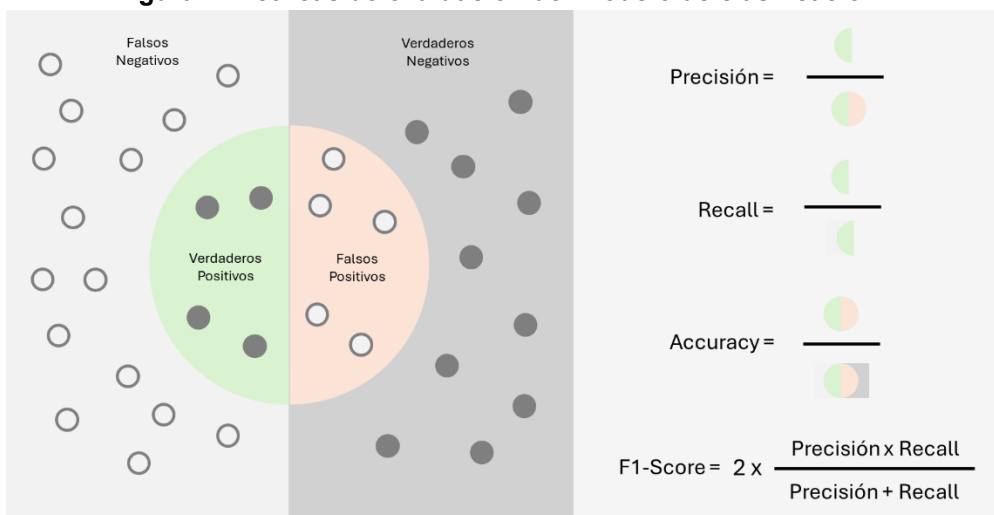
- Sensibilidad o Recall

$$Recall = \frac{VP}{VP + FN}$$

- F1-score (métrica que combina precisión y recall en un solo indicador de desempeño)

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Figura 7. Métricas de evaluación del modelo de clasificación



Estas métricas permiten evaluar no solo la exactitud global del modelo, sino también su capacidad para identificar correctamente periodos de aumento en los depósitos. Esto resulta especialmente relevante desde una perspectiva de monitoreo y gestión de liquidez, donde los costos asociados a errores de clasificación pueden ser asimétricos.

Interpretación del modelo

Si el modelo SVM muestra un desempeño predictivo superior al de un clasificador aleatorio, se interpreta como evidencia de que los indicadores de sentimiento contienen información relevante para anticipar cambios en el saldo de los depósitos. Este resultado sugiere que el análisis de sentimiento puede funcionar como una señal temprana complementaria a los indicadores financieros tradicionales.

APLICACIONES Y USO

El modelo propuesto puede ser utilizado como una herramienta complementaria para el análisis y monitoreo del comportamiento de los depósitos bancarios, integrando información alternativa proveniente de redes sociales con técnicas de Machine Learning. Su principal valor radica en la capacidad de transformar información cualitativa en señales cuantitativas que apoyen la toma de decisiones en distintos ámbitos de la gestión financiera y del riesgo.

Monitoreo temprano de cambios en los depósitos

El modelo permite el monitoreo continuo del sentimiento en redes sociales como indicador anticipado de cambios en los saldos de depósitos. La detección temprana de deterioros en el sentimiento genera alertas sobre posibles salidas de depósitos, facilitando la gestión preventiva de la liquidez.

Activación del Plan de Financiamiento de Contingencia

El modelo puede servir como un mecanismo para activar el plan de financiamiento de contingencias de las instituciones bancarias. Al proporcionar señales sobre deterioros en el sentimiento y anticipar cambios en los depósitos, permite a los bancos implementar medidas de financiamiento adecuadas en situaciones críticas.

Apoyo a la gestión del riesgo de liquidez

Este modelo se puede integrar como fuente de información adicional en los marcos de gestión del riesgo de liquidez. Proporciona señales sobre la dirección esperada de los saldos de depósitos, complementando los análisis tradicionales y regulatorios.

Evaluación del impacto Riesgo Reputacional

Los indicadores de sentimiento reflejan la percepción pública hacia el banco y pueden utilizarse para evaluar el impacto de eventos reputacionales, campañas de comunicación o cambios estratégicos. El seguimiento del sentimiento antes y después de estos eventos proporciona información valiosa para la toma de decisiones a los órganos de gobierno y alta dirección.

Uso exploratorio y analítico

El modelo también se puede emplear con fines exploratorios para analizar la relación entre el sentimiento en redes sociales y variables financieras a lo largo del tiempo. Este enfoque ayuda a identificar patrones y evaluar hipótesis, sentando las bases para el desarrollo de modelos más avanzados no solo en materia de riesgo de liquidez, sino en otras áreas de las instituciones financieras, o para la incorporación de nuevas fuentes de información alternativa en el análisis financiero.

SUPUESTOS TÉCNICOS

El modelo se basa en varios supuestos que simplifican la complejidad del fenómeno y permiten su implementación. Estos no pretenden describir exhaustivamente todos los factores que influyen en los saldos de depósitos, sino establecer un marco coherente para utilizar la información de redes sociales como señal predictiva.

- 1. Representatividad de comentarios:** Se asume que los comentarios en redes sociales son una muestra representativa del sentimiento hacia el banco, reflejando percepciones relevantes del público, especialmente en momentos de alta interacción digital.
- 2. Estabilidad del sentimiento:** Se define que la relación entre el contenido de los comentarios y la clasificación de sentimientos es estable a lo largo del tiempo, sin cambios significativos en el lenguaje o las reglas de clasificación.
- 3. Validez de los indicadores de sentimiento:** Se establece como supuesto que los indicadores construidos capturan adecuadamente el tono predominante de la conversación en redes sociales y son comparables a lo largo del tiempo.
- 4. Relación entre sentimiento y depósitos:** Se pone como punto de partida que existe una relación relevante entre los indicadores de sentimiento y la dirección de los cambios en los saldos de depósitos, entendida como predictiva.
- 5. Estabilidad temporal del modelo:** la relación aprendida por el modelo durante el entrenamiento es válida durante su evaluación y aplicación, sin cambios estructurales abruptos.
- 6. Independencia de factores externos:** los efectos de factores macroeconómicos o regulatorios no dominan la relación entre el sentimiento y los depósitos, permitiendo que los indicadores de sentimiento aporten información adicional.

RIESGOS Y LIMITANTES

El modelo presenta ciertos riesgos y limitaciones que deben considerarse al interpretar sus resultados. Estas limitaciones surgen tanto de la naturaleza de los datos como de las decisiones metodológicas.

1. **Sesgo en redes sociales:** Los comentarios pueden no representar adecuadamente a todos los depositantes, generando una sobre-representación de opiniones extremas.
2. **Ruido en el lenguaje:** El lenguaje informal y ambiguo de las redes sociales puede introducir errores en la clasificación de sentimientos, afectando el rendimiento del modelo.
3. **Limitaciones del análisis de sentimientos:** La efectividad del modelo depende de la calidad del análisis inicial. Cambios en el lenguaje o en el comportamiento de los usuarios pueden disminuir su capacidad predictiva.
4. **Riesgo de sobre-ajuste:** Existe el riesgo de que el modelo se ajuste en exceso a los datos históricos, capturando patrones que no se repiten en el tiempo.
5. **Interpretación limitada:** El modelo predice la dirección del cambio en los saldos de depósitos, pero no estima el tamaño (volumen) del efecto ni establece relaciones causales.
6. **Sensibilidad a eventos excepcionales:** Eventos extraordinarios pueden alterar el comportamiento de los depositantes y la dinámica de redes sociales, limitando la capacidad predictiva del modelo.

CONCLUSIÓN REFLEXIVA

El análisis de sentimientos aplicado a información proveniente de redes sociales constituye una herramienta complementaria relevante para anticipar cambios en el comportamiento de los depósitos bancarios. Al transformar datos no estructurados en indicadores cuantitativos, el modelo propuesto permite incorporar una dimensión adicional de percepción y riesgo reputacional al monitoreo tradicional de la liquidez. Si bien presenta limitaciones inherentes a la naturaleza de los datos y a los supuestos metodológicos, su uso como señal temprana puede fortalecer la gestión preventiva del riesgo de liquidez y apoyar la toma de decisiones en contextos de tensión financiera.

REFERENCIAS

- (1) Baker, S. R., Bloom, N., & Davis, S. J. (2016). *Measuring economic policy uncertainty*. The Quarterly Journal of Economics, 131(4), 1593–1636.
- (2) Banco de México. (2024, diciembre). *Reporte de estabilidad financiera: Diciembre 2024*. Banco de México.
- (3) Bank for International Settlements. (2022). *Project Ellipse: Regulatory reporting and data analytics platform*. BIS Innovation Hub.
- (4) Cookson, J. A., Fox, C., Gil-Bazo, J., Imbet, J. F., & Schiller, C. (2023). *Social media as a bank run catalyst*. FDIC Bank Research Conference.
- (5) Del Sarto, N., Bocchialini, E., Gai, L., & Ielasi, F. (2024). *Digital banking: how social media is shaping the game*. Qualitative Research in Financial Markets.
- (6) Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)* (pp. 4171–4186). Association for Computational Linguistics.
- (7) Giuliana, R., Panfilo, M., & Peltonen, T. (2024). *Deposit flows during monetary tightening: The role of digital banking and social media*. European Central Bank / ESRB Workshop.
- (8) Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- (9) Jurafsky, D., & Martin, J. H. (2026). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed., draft). Online manuscript.
- (10) Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- (11) Loughran, T., & McDonald, B. (2011). *When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks*. Journal of Finance, 66(1), 35–65.
- (12) Pang, B., Lee, L., & Vaithyanathan, S. (2002). *Thumbs up? Sentiment classification using Machine Learning techniques*. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 79–86). Association for Computational Linguistics.
- (13) Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). *Lexicon-based methods for sentiment analysis*. Computational Linguistics, 37(2), 267–307.
- (14) Tetlock, P. C. (2007). *Giving content to investor sentiment: The role of media in the stock market*. The Journal of Finance, 62(3), 1139–1168.

DISCLAIMER

Este documento ha sido preparado por Grupo Financiero Banorte, S.A.B. de C.V. ("Banorte") para fines meramente informativos, utilizando fuentes públicas y especializadas consideradas confiables; no obstante, Banorte no garantiza la precisión, integridad, ni la vigencia de la información prevista en el mismo. Su contenido no constituye asesoría legal, fiscal, financiera, contable ni una interpretación oficial del marco legal aplicable. En caso de requerirlo, se recomienda consultar con asesores legales, fiscales, financieros, contables o de inversión independientes. La información contenida en este documento está sujeta a modificaciones sin previo aviso.

Ni Banorte ni ninguna de las entidades que integran el Grupo serán responsables, en ningún caso, por pérdidas, daños o perjuicios que pudieran derivarse, directa o indirectamente, del uso de este documento o de su contenido. Del mismo modo, Banorte no adquiere compromiso alguno de actualizar la información aquí contenida ni de notificar cambios posteriores. El contenido de este documento podría diferir de la opinión o interpretación de autoridades financieras nacionales o internacionales, y no debe considerarse como un posicionamiento institucional de Banorte.

RIESGOS FINANCIEROS

