

Un puente digital hacia la inversión: Democratizando las finanzas avanzadas frente a la crisis de coste de vida

Carles Villar Vegara

RESUMEN:

Este trabajo aborda la problemática de la educación financiera en la juventud española, un colectivo marcado por un desconocimiento financiero marcado que deriva en inestabilidad económica y dificulta el ahorro. Basándose en datos de *El Economista*, se identifica que un 62% de los jóvenes no invierte por desconocimiento técnico o por la creencia de que se necesita un gran capital para empezar a invertir.

La investigación constituye la base empírica para el desarrollo de la aplicación “Horizonte”, proyecto de mi Trabajo de Fin de Grado, cuyo objetivo es democratizar el acceso a la planificación financiera. Para ello, se utiliza como fuente de datos las comunidades de reddit, analizando los datos no estructurados que permiten capturar miedos y aspiraciones, con el fin de entender que posibles herramientas pueden ser útiles para crear una herramienta que ayude realmente a los ciudadanos.

Los métodos que usaremos serán los vistos en clase como técnicas de procesamiento de lenguaje natural (NLP) y modelado de tópicos (LDA) en r, para este informe se cuenta con más de 4000 testimonios para diferenciar las preocupaciones. Los resultados permiten justificar técnicamente el diseño de funcionalidades específicas que puedan ayudar a mejorar la situación actual. En conclusión, mediante este estudio vamos a transformar los datos de las redes sociales en información para garantizar que Horizonte responde a problemas reales de los usuarios y fomentar la autonomía financiera segura en un entorno económico complejo.

INTRODUCCIÓN:

El siguiente trabajo nace con la motivación de entender un problema estructural en la juventud actual: la profunda carencia de educación financiera y la dificultad de gestionar las finanzas en un entorno económico cada vez más complejo, que deriva en problemas de ahorro, de vivienda, jubilación y calidad de vida en general.

La relevancia de este proyecto se hace evidente al observar que la barrera de entrada no es solo falta de capital, sino también de conocimiento. Según datos publicados por **El Economista (2025)**, “un 62% de los que actualmente no invierten porque no saben cómo hacerlo, cuáles son los beneficios, en qué invertir o cómo empezar”, estos datos definen el punto de dolor de la juventud española, una parálisis provocada por la falta de una hoja de ruta o guía.

Motivado por ofrecer una solución real, este trabajo constituye la base empírica para mi trabajo de fin de grado, centrado en el desarrollo de “Horizonte”. El objetivo de Horizonte es ser una herramienta de apoyo financiero accesible para todo el mundo que ayude a la gente a gestionar y conseguir una estabilidad financiera que cada vez es más difícil en España, una forma de democratizar el acceso a la planificación financiera.

Por ello este trabajo es muy útil y se alinea muy bien con mi TFG ya que me va permitir detectar mediante la evidencia real, las funcionalidades reales que resuelvan dudas y problemas que la gente realmente tiene y no basarme en mi intuición.

Para garantizar que “Horizonte” responde a necesidades reales, se ha seleccionado la comunidad de Reddit como fuente de información. El objetivo es analizar cuáles son los miedos, problemas y aspiraciones, para a partir de esto saber que herramientas pueden ser útiles.

Los objetivos específicos de este trabajo son: identificar los temas que generan mayor preocupación (vivienda, inversión, fiscalidad, etc.), medir la intensidad emocional como el miedo incertidumbre o confianza hacia el futuro financiero, validar la viabilidad e funcionalidades técnicas para conocer qué funcionalidades responden a demandas reales expresadas por los usuarios.

MARCO TEÓRICO:

2.1. Definición y Realidad de la Alfabetización Financiera

La educación financiera no es solo saber qué es el dinero o qué es un banco, es la capacidad de tomar decisiones a partir de información económica de forma inteligente y que afecten a nuestro bolsillo. Existen diversos estudios sobre la situación actual de la educación financiera y muy pocos muestran resultados favorables. Según la Encuesta de Competencias Financieras (ECF) elaborada por el Banco de España y analizada por Funcas, donde se preguntan las famosas “big three” que se refiere a 3 conceptos clave que se usan para medir el conocimiento financiero básico de una persona, Interés compuesto, inflación y diversificación de riesgo. Según este estudio, apenas el 19% de los españoles es capaz de responder correctamente a las 3 preguntas básicas.

2.2. El Sesgo del Exceso de Confianza y el Riesgo de Inflación

Este dato es demoledor ya que significa que 8 de cada 10 personas no tienen conocimientos mínimos para gestionar sus ahorros sin perder dinero. Además, el problema se agrava con lo que señala una noticia de *El País* sobre un estudio realizado por el banco santander que alerta sobre un problema de exceso de confianza: “La mayoría de los españoles cree que domina sus finanzas, pero solo uno de cada 3 entiende realmente cómo funciona la inflación”. Esto demuestra que mucha gente está perdiendo dinero sin saberlo, simplemente creen que saben más de lo que realmente entienden, lo cual es un peligro real para sus ahorros a largo plazo. No estamos hablando de ser expertos inversores y poder vivir de la inversión, pero sí de entender conceptos básicos que afectan a nuestros ahorros y saber cómo tratarlos para no perder capacidad adquisitiva a largo plazo.

2.3. Barreras Psicológicas y el Miedo a la Inversión

Más allá del desconocimiento técnico, existen barreras psicológicas que frenan el ahorro. Según datos de *El Asesor Financiero*, seis de cada diez españoles que no invierten, lo hacen por falta de educación financiera. Este miedo nacido de la inseguridad, provoca que el pequeño ahorrador se mantenga al margen de los mercados perdiendo poder adquisitivo, por el simple hecho de pensar que para invertir tienes que tener un alto conocimiento del mercado, sin tener en cuenta que existen otras estrategias que sirven para ahorrar que son más simples, sin falta de conocimientos excesivos y que dan mejores resultados que el ahorro tradicional sin dolores de cabeza.

2.4. El Auge de los "Finfluencers" y la Crisis de Confianza Institucional

Como consecuencia de este miedo y de las creencias de que el asesoramiento profesional es inaccesible porque es para gente rica o por inseguridad, ha surgido un fenómeno peligroso llamado "Finfluencers". Según reporta *El país* (2024), los miembros de la Generación Z sustituyen a los asesores expertos por "gurús" de redes sociales que prometen rentabilidades irreales sin ningún tipo de garantía sobre sus conocimientos.

La combinación de estos factores, desconocimiento, exceso de confianza y percepción del asesoramiento elitista, genera una situación crítica. Este vacío de confianza en las instituciones tradicionales ha permitido el auge de los "finfluencers". Estos perfiles aprovechan la facilidad de acceso y cercanía digital para ofrecer consejos sin rigor técnico. El problema es que el usuario, motivado por la inseguridad y falta de alternativas asequibles, deposita en ellos una confianza que a menudo se transforma en decisiones erróneas.

Ante este escenario, se plantea la cuestión ¿Existe alguna solución tecnológica capaz de democratizar el acceso a la planificación financiera sin barreras de coste y lenguaje de banca tradicional? y, sobre todo, ¿Cómo podemos identificar, a través del análisis de datos reales, las funcionalidades que responden a las necesidades reales de los usuarios?

METODOLOGÍA Y DATOS:

3.1. Diseño de la Investigación y Objetivos Metodológicos

A continuación se detalla la estructura metodológica que se ha seguido para realizar este estudio, con el objetivo central de transformar el discurso digital y no estructurado en información relevante sobre la situación de la juventud, con idea de obtener cuales son los problemas que más se repiten e intentar encontrar una solución a estos a través de la integración de técnicas de procesamiento de lenguaje natural se buscan 2 cosas.

Desentrañar la estructura latente, superar la simple observación de palabras clave para identificar, mediante el modelo LDA con $k = 6$, los núcleos temáticos reales que configuran la agenda de preocupaciones financieras.

En segundo lugar cuantificar la carga afectiva, se ha conseguido descifrar la intensidad emocional del discurso, permitiendo entender si los temas detectados generan reacciones de ira, miedo o confianza.

3.2. Fuente de Datos: La Elección de Reddit

La fuente principal de información ha sido la plataforma de Reddit, específicamente seleccionando comunidades donde predominan los perfiles jóvenes y el discurso gira en torno a la economía, finanzas, empleo y vivienda.

La elección de Reddit como fuente principal de datos para este estudio, responde a la naturaleza misma del objetivo de estudio: las preocupaciones financieras de los jóvenes. A diferencia de otras redes sociales más centradas en la proyección de una imagen perfecta como Instagram o una validación profesional como LinkedIn, Reddit se basa en el pseudónimo lo que permite encontrar opiniones y relatos más reales y sinceros sin la necesidad de mostrar una imagen idealizada que realmente no existe.

Esta característica fomenta un espacio donde los usuarios comparten detalles específicos sobre sus frustraciones y miedos económicos sin la mancha social que conlleva el fracaso financiero en la vida real, lo que permite obtener una densidad semántica superior, con comentarios más extensos, descriptivos que permiten una captura más rica de la realidad socioeconómica .

Otra razón de peso para usar Reddit y no X u otra red social, se justifica con la viabilidad técnica en el acceso a los datos. Mientras que otras redes han restringido sus APIs o han implementado métodos de pago para la extracción de datos, Reddit mantiene una infraestructura que permita la extracción de información a través de su API oficial.

3.3. Procedimiento de Extracción y Selección de la Muestra

Para la extracción se ha usado la API de Reddit para realizar el scraping selectivo. El proceso se concentra en hilos de discusión con palabras específicas como “invertir”, “jubilación”, “pensión”, “ahorro”, “empleo”, etc en diferentes subreddits. Para evitar un sesgo de clase y contar con una muestra con opiniones de perfiles de usuario variados, se ha realizado una extracción de datos entre diferentes comunidades de reddit, como lo son r/SpainFire donde encontramos un perfil más orientado a la inversión, r/es que es una comunidad nacional más general y r/askspain donde también encontramos consultas de carácter cotidiano. Esto nos permite una captura más amplia de varios perfiles de usuarios.

Dada la naturaleza transversal de los conceptos buscados, muchos hilos de discusión estaban duplicados, por eso implementamos un proceso de deduplicación basado en las URL, asegurando que cada hilo fuera analizado una sola vez, con el objetivo de optimizar la muestra y eliminar el sesgo de repetición y agrupamos por comunidad.

Finalmente para garantizar la calidad de la muestra y que el análisis se basará en hilos de alta densidad de debate, se aplicaron unos filtros. El primer filtro fue eliminar todos los hilos que tuviesen menos de 10 comentarios, descartando publicaciones con contenido superficial. Esta decisión permite centrar el análisis en los debates más influyentes y representativos garantizando la viabilidad computacional.

El siguiente paso importante ha sido el proceso de limpieza donde nos hemos encargado de eliminar el ruido sistémico como las etiquetas automáticas de la plataforma como [delete] o [removed], artefactos visuales que hacen referencia a archivos multimedia como .jpeg o .png, entidades no semánticas como URLs, menciones a usuarios o subreddits que no aportan valor y por último la normalización, eliminando espacios y caracteres especiales preservando únicamente signos de puntuación que pueden tener una carga emocional que nos interesa para el análisis como exclamaciones e interrogaciones.

3.4. Procesamiento, Limpieza y Normalización del Corpus

El corpus final tras el proceso de limpieza y el filtrado de hilos con baja interacción, se compone de 4229 comentarios. Este volumen garantiza una densidad estadística suficiente para que el modelado de temas LDA identifique patrones.

Las Stopwords que se han utilizado para el análisis se han sacado de un repositorio de github, la ventaja de estas stopwords es que son una lista más completa que las que se pueden conseguir mediante librerías de r, en un corpus tan informal y tan grande como con el que estamos trabajando es importante tener una lista completa de stopwords para no acabar con un análisis vacío, además después de ver los resultados se han encontrado aun más palabras que generaban ruido y no aportan

información de interés como año, años, gracias o algunas palabras en inglés por lo que se han añadido a la lista de stopwords para conseguir unos mejores resultados.

3.5. Representación de Datos

En cuanto a las representaciones trabajadas, para la ejecución de las técnicas analíticas, se han transformado los datos en dos formatos.

El primero ha sido el formato Tidy, en este caso se ha dividido el texto en tokens para permitir la integración con el diccionario de sentimientos. Esta representación facilita el cálculo de la intensidad emocional.

El segundo formato, la Matriz Documento-Término, se ha construido una matriz esperanza de frecuencias necesaria para el modelado probabilístico de tópicos. En esta estructura, cada fila representa un comentario y cada columna un término del vocabulario, esto permite al algoritmo de LDA que usaremos identificar patrones en los datos.

Para el proceso analítico del corpus, se han seleccionado dos técnicas que operan en diferentes niveles: el análisis de sentimiento basado en léxicos y el modelado probabilístico de tópicos.

En primer lugar para el análisis de sentimiento y la detección de emociones se ha optado por un enfoque léxico-basado utilizando el diccionario NRC a través de la librería syuzhet. Esta elección se justifica por dos grandes razones.

A diferencia de modelos binarios que solo ofrecen polaridad positiva o negativo, con NRC ha permitido extraer ocho emociones básicas que se representan en el corpus como ira, anticipación, asco, miedo, alegría, tristeza, sorpresa y confianza.

Esto es crucial para entender no solo si los jóvenes están descontentos, sino qué tipo de sentimiento domina su discurso financiero.

El segundo motivo importante por el que se ha decantado por este léxico es la adaptación al castellano, el corpus que estamos trabajando está en castellano ya que queremos estudiar el mercado español, esto significa una complejidad morfosintáctica superior al inglés, especialmente en el uso de negaciones y modificadores. Además la mayoría de diccionarios están adaptados y optimizados para proporcionar un buen rendimiento en inglés y esta opción permite el uso de un diccionario NRC en español que nos permite una cobertura léxica amplia.

La segunda técnica que se ha aplicado para el análisis del corpus es el modelado de tópicos o LDA, con la que queremos extraer información importante de forma automática, esta técnica es la mejor opción por varios motivos.

Primeramente por la detección de patrones latentes, al ser un algoritmo no supervisado, el LDA nos permite identificar temas basándose en patrones de co-ocurrencia estadística de las palabras. Con un corpus de 4000 comentarios, donde el análisis manual es inviable, el LDA permite descubrir de qué se habla realmente en las comunidades de nuestro corpus.

En segundo lugar, este método permite un análisis multi tópico, es decir el modelo asume que cada comentario puede contener diversos temas. Lo que nos proporciona una mayor fiabilidad en un

entorno como Reddit donde el usuario puede hablar de diferentes temas en el mismo comentario como “sueldo” e “inflación”.

Finalmente esta técnica nos ha proporcionado una visión más completa de la variedad semántica del corpus, con el LDA se ha conseguido organizar el lenguaje en dimensiones semánticas permitiendo transformar tokens aislados en conceptos accesibles y relevantes para el objetivo del trabajo.

Para la determinación del valor de k , se ha usado un proceso de optimización del hiper parámetro K , con el cual se ha determinado cuál es el valor de k que mejor explica la estructura latente del corpus, siguiendo la norma el punto donde se minimiza CaoJuan2009 y maximiza Deveaud2014 es en $k = 5$ por lo que en el análisis se ha evaluado el valor de $k = 5$ y $k = 6$ para determinar cuál era el más óptimo.

Tras la evaluación de ambos, se determinó que el modelo de 6 tópicos ofrece el nivel más óptimo de granularidad semántica para el corpus. Un $k = 5$ presentaba una cohesión temática insuficiente fusionando categorías mientras que un $k = 6$ logra aislar las dimensiones clave que encontramos que son inflación y consumo, gestión de gastos, mercado inmobiliario, planificación de futuro, contexto Político-Fiscal y financiación e Hipotecas.

3.6. Resolución de Conflictos Técnicos: Trazabilidad e IDs

El único problema técnico significativo se encontró en la gestión de las IDs. Debido a que la estructura de Reddit es jerárquica y ramificada, con códigos que indican la relación entre comentarios y respuestas, se generó un conflicto de sincronización al transitar entre diferentes representaciones de los datos.

Este problema se manifestó mediante la aparición de valores nulos, la discrepancia entre los IDs jerárquicos de la fuente original y los índices generados por las funciones de R impedía la unión de los resultados, lo que se traduce en una pérdida de la trazabilidad en gran parte de la muestra.

La solución aplicada fue una reindexación numérica de las 4000 observaciones tras la fase de limpieza. Al asignar un ID plano y secuencial, se garantiza una correspondencia uno a uno entre el análisis de sentimiento NRC y la clasificación de temática LDA, asegurando la integridad total de la base de datos final.

El último y más importante de los pasos ha sido la representación de la información obtenida que se ha realizado mediante el uso de gráficos que veremos en el siguiente apartado para sacar las conclusiones.

RESULTADOS :

4.1. Identificación de Dimensiones Críticas (LDA)

Como hemos mencionado en anteriores apartados el objetivo del análisis es encontrar los problemas y preocupaciones más comunes y en base a estos proponer una solución digital que ayude a los ciudadanos a conseguir la independencia y tranquilidad financiera. Como ya sabemos la situación actual presenta muchas lagunas para los ciudadanos en cuanto a la vivienda y la gestión del coste de vida, pero los datos de Reddit nos permiten ponerle nombre y apellido a estas deficiencias.

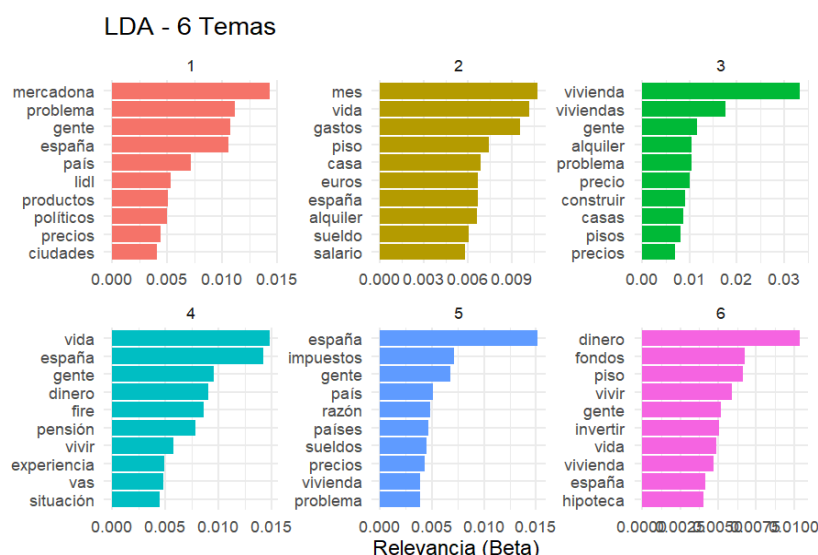


figura 1

Como podemos observar en la figura 1, las preocupaciones detectadas se dividen en estas 6 grandes dimensiones. Este gráfico de relevancia nos permite visualizar qué términos definen cada conversación latente del corpus.

La primera temática que encontramos se centra en la microeconomía. La prominencia de términos como “Mercadona”, “lidl” y “precios” revela que la inflación es un concepto que afecta al usuario y lo ve como una barrera cotidiana con la subida de precios de productos de primera necesidad como lo son alimentos u otros que se venden en los supermercados. El discurso gira en torno al encarecimiento de la cesta, actuando como el primer nivel de estrés financiero que impide cualquier capacidad de ahorro.

En la segunda temática encontramos términos como “sueldo”, “gastos” y “mes”, lo cual se puede interpretar como una narrativa de supervivencia donde los ingresos se consumen por los gastos y cuesta llegar a final de mes. En este contexto, el “sueldo” no se percibe como una herramienta de progreso, sino como un recurso limitado que apenas garantiza una estabilidad económica mensual y menos aún permite el ahorro de los ciudadanos, para conseguir un colchón de seguridad que les permita pagar posibles gastos de emergencia como avería de coche y permita una tranquilidad mayor.

El tercer tema está compuesto por palabras como “viviendas”, “problema” y “construir”, esto indica una preocupación por la falta de oferta y el estado del mercado inmobiliario. Los jóvenes discuten la vivienda como un bien inaccesible, analizando por qué no se construye o por qué los precios han perdido total relación con la realidad salarial.

La temática cuatro representa la respuesta proactiva, términos como “pensión”, “fire” e “dinero” muestran un perfil de usuario que ante la desconfianza en los sistemas públicos busca construir su propia seguridad financiera mediante estrategias activas. Esta es la dimensión con mayor carga de planificación y donde la educación financiera aparece como la única vía de escape percibida. Términos como fire que es el acrónimo de financial independence, retire early, es una filosofía de vida y estrategia de inversión que ha ganado fuerza entre los jóvenes que buscan escapar del ciclo de trabajo tradicional hasta los 65-67 años.

Siguiendo por la temática 5 encontramos términos como “problema”, “impuestos” y “país”, esta dimensión enmarca la situación económica como una consecuencia de la gestión estatal. El discurso sugiere que el sistema fiscal es percibido como una carga adicional que castiga al ahorro y dificulta la libertad financiera de muchos ciudadanos, alimentando un sentimiento de rechazo hacia la estructura institucional actual.

Finalmente en la temática 6 analiza la entrada al mercado de activos, que a diferencia de la temática 3 se centra en la viabilidad financiera, encontramos términos como “hipoteca”, “fondos” y “dinero”. Aquí podemos encontrar la frustración del que tiene ingresos pero no suficiente ahorro para una entrada o del que busca productos de inversión pero no sabe cómo gestionarlo o se encuentra con un sistema bancario que percibe como hostil.

4.2. Análisis de sentimiento medio por tópico

Una vez sabemos cuales son las principales dimensiones financieras que preocupan al ciudadano español, vamos a ver cómo se clasifican según el sentimiento medio hacia cada uno.

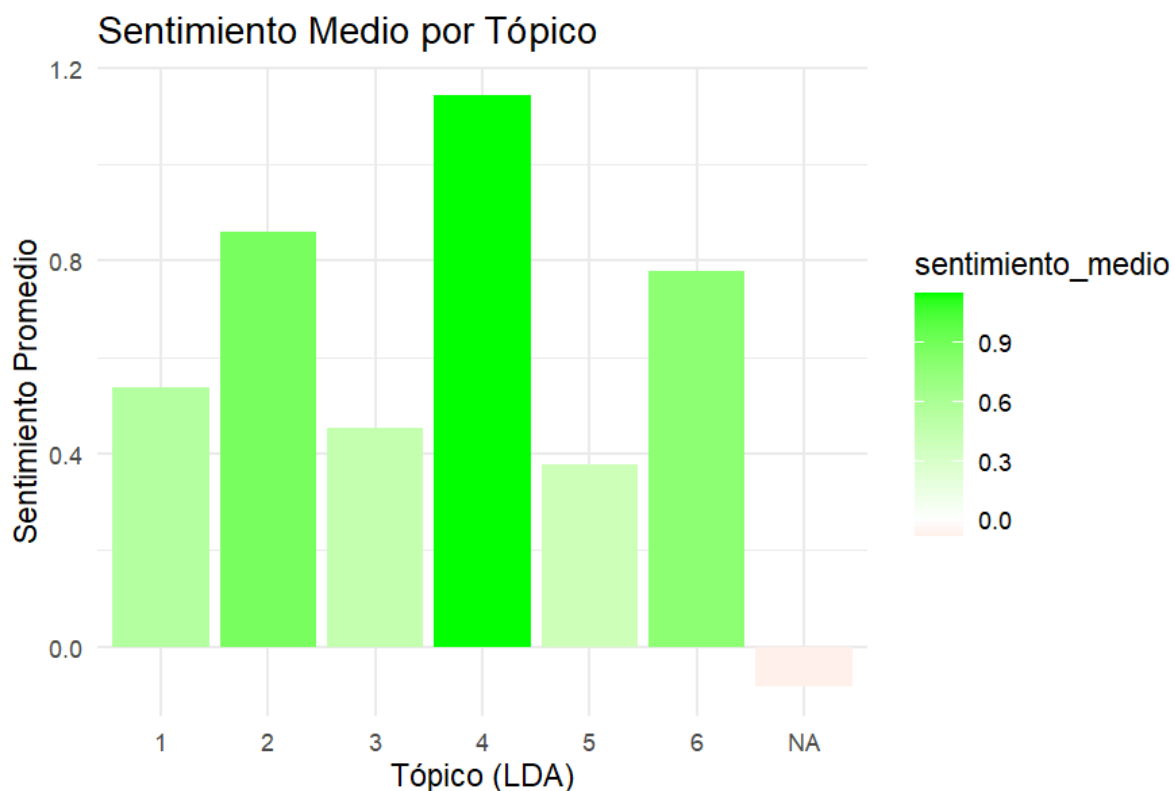


figura 2

Lo primero que destaca al ver el gráfico es el tópico 4, que registra el sentimiento promedio más elevado de toda la muestra, superando el valor de 1,1. Como hemos mencionado este tópico está dominado por conceptos como “vida”, “fire” y “pensión”, lo que hace evidente que este optimismo no sea casual. Representa la narrativa de la autonomía financiera, es el único tópico que no se centra en la queja si no en la planificación de su futuro mediante estrategias e inversión.

En el extremo opuesto encontramos el tópico 5, que presenta el nivel de sentimiento más bajo cerca de 0,35. la baja puntuación en este apartado confirma la desafección institucional, el ciudadano percibe el

marco fiscal y político de España no como un apoyo, sino como un obstáculo directo que frena la capacidad de progreso económico.

Por otro lado existe contraste entre el tópico 3 y 6, ambos relacionados con la vivienda. El tópico 3, centrado en la vertiente estructural y el problema del alquiler muestra un sentimiento significativamente menos, con un 0,45 respecto a un 0,77. Esta diferencia sugiere que, mientras que el discurso del acceso a la vivienda y los precios de alquiler se viven con negatividad, el discurso sobre fondos, hipotecas e invertir mantiene una carga más aspiracional. Esto se puede explicar por la diferencia de perfiles que comentan en cada temática, mientras que en la temática tres podemos encontrar un perfil más humilde que tiene un salario que le de justo para pasar el mes, la temática 6 puede predominar un perfil de inversores con una estabilidad económica más amplia que discuten por la compra de bienes inmuebles como inversiones lo cual se ve con más optimismo al ser un mercado tan saturado y con tanta demanda y tan poca oferta.

Finalmente el tópico 1 se sitúa en una zona más moderada con un sentimiento de 0,53, esta temática describe cómo se siente el ciudadano promedio respecto a la pérdida de poder adquisitivo, observamos palabras como políticos mercadona y sugiere una visión marcadamente crítica hacia las políticas económicas actuales en España, a las que se responsabiliza directamente del aumento de la inflación.

Sin embargo, el análisis revela una brecha de alfabetización financiera significativa, el ciudadano tiende a percibir la inflación como un fenómeno externo o una decisión corporativa de las grandes empresas, cuando en realidad se trata de la pérdida de valor de la moneda fiat. Esta incomprensión de las políticas monetarias que aceleran el proceso de empobrecimiento genera un sentimiento de negatividad, el usuario detecta que su dinero vale menos, pero no llega a comprender la raíz del problema, su respuesta se limita a la queja contra los mercados.

Cabe destacar que este valor más moderado del tópico 1 se puede explicar por la alfabetización de los usuarios que comentan en estos hilos, estos usuarios que comentan en estos foros de reddit poseen una alfabetización superior a la de los ciudadanos por lo puede alterar la métrica de el sentimiento observada al igual que en otros tópicos y que no sea completamente representativo de la población por el sesgo de autoselección, aunque sí que puede dar una visión amplia de lo que piensan muchos ciudadanos.

4.3. Intensidad Emocional: El Impacto del Miedo y la Vulnerabilidad

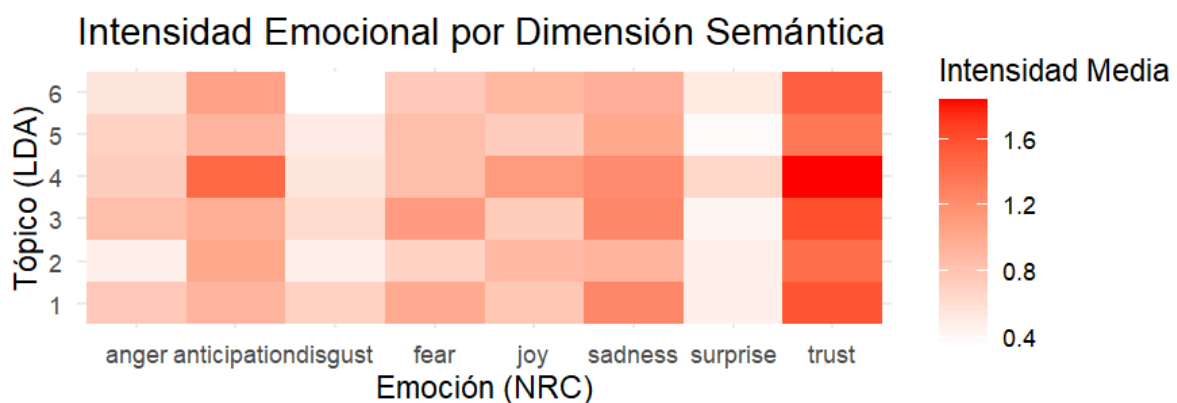


figura 3

En cuanto a las emociones, más allá de una visión positiva o negativa podemos realizar un análisis de qué emoción es la que predomina en cada temática dentro de las 8 emociones del diccionario NRC.

En el tópicos 4 observamos que predominan la confianza y la anticipación, es decir el perfil de este grupo no se traduce solo en palabras positivas, sino en una mirada puesta en el futuro. La alta intensidad en anticipación confirma que el movimiento FIRE es un proyecto de vida con una visión a largo plazo.

En cuanto a las temáticas 1 y 3 podemos observar que concentran niveles intensos de tristeza y miedo. Esta convergencia emocional no es circunstancial. Responde a una sensación de vulnerabilidad frente al encarecimiento de la vida.

Mientras otros temas pueden generar indignación, el discurso sobre el precio de productos de primera necesidad y el acceso a la vivienda está marcado por una narrativa de pérdida. Los términos predominantes como “precio”, “alquiler” y “problema” actúan como disparadores de tristeza, fruto de la impotencia del ciudadano al ver que su capacidad de subsistencia se contrae de forma constante.

Además observamos la presencia de un miedo marcado en estos tópicos que revela que el descontento no solo se refiere al presente, sino a la inseguridad hacia el futuro. En estas circunstancias el miedo actúa como una emoción paralizante. Los usuarios de Reddit expresan miedo a que la situación lejos de estabilizarse, continúe empeorando hacia la imposibilidad de pagar el alquiler o la cobertura de necesidades básicas. Este bloqueo emocional es una de las lagunas más graves detectadas, ya que la imposibilidad de ahorro dificulta la planificación financiera.

Una observación técnica de este gráfico es que la emoción confianza es intensa en todos los tópicos. Esto no significa que los comentarios sean prósperos y optimistas, sino que el diccionario NRC, detecta palabras como casa, dinero banco o padres y las categoriza como confianza aunque en el contexto de la oración no sea así, por lo que más que confianza real en la situación actual, este dato refleja la naturaleza familiar de los temas tratados.

4.4. Análisis de Bigramas: Supervivencia vs. Estrategia

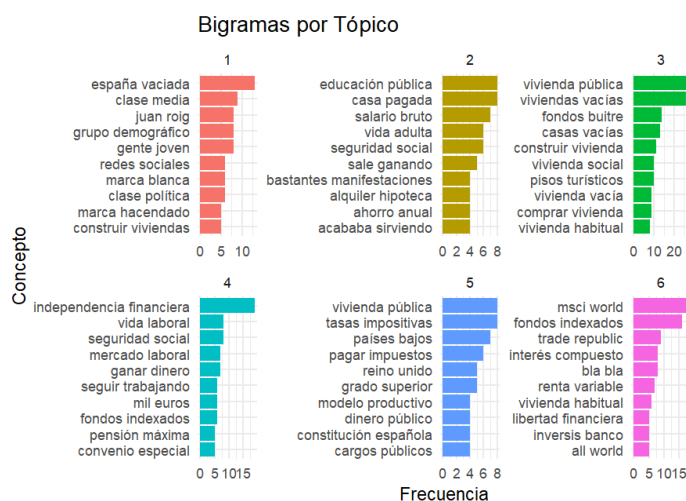


figura 4

Tras analizar las temáticas más preocupantes de los ciudadanos, cómo se sienten respecto a estas preocupaciones y la intensidad de sus emociones, vamos a terminar con un análisis de bigramas para extraer información aún más profunda de cada dimensión. Con este enfoque se pretende aterrizar las preocupaciones abstractas en conceptos operativos concretos. Al observar cómo se relacionan los términos entre sí, podemos identificar relaciones semánticas exactas que definen la inquietud financiera y percepción de vulnerabilidad de los ciudadanos.

Al observar las relaciones de palabras con mayor información analítica, destaca una división entre el lenguaje de supervivencia y el de estrategia. En el bloque de consumo y vivienda (Tópico 1 y 3) encontramos relaciones de palabras como “marca blanca”, “viviendas vacías”, “pisos turísticos” que nos proporcionan una percepción del abuso y falta de recursos básicos. Estos términos no solo describen el coste de vida, sino que revelan una mentalidad de resistencia frente a la oferta inmobiliaria especulativa y encarecimiento de la cesta de la compra básica que obliga a refugiarse en marcas de bajo coste.

En un sentimiento opuesto encontramos tópicos estratégicos que revelan una narrativa defensiva frente a la pérdida de poder adquisitivo. Las combinaciones presentes en el Tópico 6, como “msci world”, “fondos indexados” e “intereses compuestos”, junto con la búsqueda de “independencia financiera” en el Tópico 4, muestran la existencia de un perfil de usuario que intenta resistirse al sistema. Para este grupo de gente la solución no pasa por la política, sino por la educación financiera y el uso de plataformas como neo bancos y brokers para gestionar su propio futuro y conseguir la estabilidad financiera.

Finalmente, la relación entre “pagar impuestos”, “tasas impositivas” y “cargos públicos” en el Tópico 5 termina de dibujar el mapa de la desconfianza institucional. Esta idea termina de reforzar la idea de que el ciudadano percibe el sistema fiscal no como una contribución a su bienestar, si no como un lastre hacia su ahorro, donde cada vez es más la presión que ejerce.

A partir del análisis detallado del contenido del corpus, podemos identificar tres núcleos de conflicto que definen el escenario financiero. En primer lugar se observa una asfixia financiera hacia la clase media y la gente joven, donde la inflación cada vez es más notoria en el día a día y se tiene que combatir mediante la compra de productos más baratos como la marca blanca. Esto confirma que el miedo y la tristeza detectados, tienen una raíz material, el encarecimiento y la pérdida de poder adquisitivo de la clase media donde cada vez se sienten más vulnerables.

En segundo lugar, existe una brecha de acceso habitacional donde el discurso sobre la “vivienda pública” y el rechazo a los “fondos buitres” evidencia una frustración estructural que el ciudadano no sabe como resolver generando un bloqueo que frena la planificación a largo plazo.

Finalmente, la desconexión con el estado, que aunque su función principal debería ser la redistribución de la riqueza y la defensa del ciudadano, parece que los usuarios cada vez lo ven más al contrario, con una presión fiscal asfixiante que no les permite ahorrar.

Esta percepción del estado tiene consecuencias directas en el relevo generacional y la evolución del país. Los datos reflejan que la gente joven se encuentra atrapada en una situación que imposibilita su progreso, por un lado la inaccesibilidad a la vivienda que impide su emancipación y por otro la

fiscalidad que drena sus ingresos. Los usuarios no solo se quejan de la situación si no que muestran una profunda desconfianza hacia la capacidad financiera.

Finalmente encontramos que para combatir estos problemas surge un movimiento de alfabetización financiera que busca refugio en soluciones privadas y estrategias individuales que les permita protegerse ya que sienten que el marco institucional ha dejado de velar por sus intereses.

DISCUSIÓN Y CONCLUSIONES:

5.1. Interpretación de resultados y diagnóstico

El análisis de los datos revela una fractura evidente entre la realidad económica percibida y las herramientas de las que dispone el ciudadano para protegerse. Como bien se ha mencionado en el inicio del trabajo, el objetivo principal del análisis es comprender problemas específicos de los ciudadanos y crear un concepto de aplicación que ayude a mejorar la situación en base a los problemas detectados.

El problema el miedo ante la situación por la vivienda y el encarecimiento de la compra es claro, pero ¿Que podemos ofrecer a la población que les ayude a conseguir una estabilidad financiera básica? Para responder a esta cuestión, la solución digital propuesta debe actuar como un catalizador que transforma la preocupación en una estrategia de defensa patrimonial, centrándose en tres pilares fundamentales que se identifican en las lagunas detectadas en los datos.

5.2. Implicaciones prácticas: La propuesta de valor de "Horizonte"

La herramienta debe actuar como un mentor digital, que traduce el esfuerzo de ahorro cotidiano en capacidad de inversión. Si el usuario utiliza estrategias de ahorro como comprar marca blanca o reducir el consumo de algunos productos, la aplicación le mostrará de forma visual como ese pequeño ahorro, aplicado a instrumentos del tópico 6 como “MSCI World” o “Interés compuesto”, pueden generar un patrimonio real a largo plazo. El objetivo es crear un sistema de educación financiera que no se vea como algo teórico sino como algo tangible que transforme la tristeza en anticipación de una cartera creciente.

También podemos crear un arquitecto de metas, con el cual el usuario pueda planificar de forma inteligente el acceso a la vivienda o la compra de un automóvil o el objetivo que tenga el usuario en x momento, de esta manera se busca educar al usuario en la creación de fondos de inversión para la consecución de sus objetivos. Proyectando las probabilidades de conseguir sus objetivos en x años en base a métricas que nos proporcione el propio usuario y el objetivo de este.

Otra herramienta útil sería un dashboard de soberanía fiscal, ante la presión generada por el sistema fiscal, se podría ofrecer un módulo de optimización tributaria legal y fiscal transparente. Educación financiera se refiere a enseñar al ciudadano cómo funcionan las tasas impositivas y qué vehículos de inversión como planes de pensiones o deducciones pueden reducir su carga impositiva. Este enfoque tiene como objetivo devolver al usuario la sensación de control y transparencia.

Otra herramienta que puede ayudar a la tranquilidad del ciudadano es un simulador de escenario de crisis, para aprender cómo reducir el riesgo ante una posible crisis. Utilizando datos históricos de inflación, mostrar de forma visual cómo se comportaría sus ahorros en diferentes escenarios y enseñar estrategias para entender como combatirlo.

Finalmente se podría plantear la idea de un sistema de educación financiera interactivo que siga estrategias prácticas y visuales para que el usuario aprenda conceptos básicos de finanzas y de esta manera pueda tomar decisiones con una base más consolidada lo que ayudaría a tomar decisiones más acertadas. También sería interesante una calculadora de independencia, en lugar de mirar cuánto tienes en el banco que te ofrezca la información de forma más atractiva como por ejemplo cuántos meses de trabajo lleva ahorrados en su fondo de inversión o que probabilidad tiene de llegar a x cantidad que seleccione el usuario en x años si sigue con la estrategia que está siguiendo y en base al ritmo que eso lo calcularía la app.

Como bien se ha mencionado anteriormente el objetivo principal es crear una app que ayude al usuario a gestionar sus finanzas de una forma activa a la vez que crea una base de conocimiento más amplia y básica de una forma interactiva y atractiva que motive al usuario a continuar con el uso de la herramienta y tal vez en un futuro a reducir la dependencia de las pensiones en el sistema.

5.3. Limitaciones del estudio

Pero todo este análisis tiene varias limitaciones, aunque nos da una visión amplia de la situación actual y podemos sacar conclusiones bastante potentes. La muestra está sesgada, nuestra muestra tiene el sesgo de autoselección, la muestra obtenida de Reddit representa a un usuario con una alfabetización financiera superior a la media, lo que sugiere que el malestar en la población podría ser incluso más profundo, como hemos visto en estos foros se habla tanto de problemas de vivienda como de estrategias para mitigar la pérdida de capacidad adquisitiva lo que sugiere que los usuarios tienen un nivel de educación financiero medio que hace que en muchos casos no represente realmente los problemas de los ciudadanos.

Otra limitación que podemos encontrar es en el diccionario NRC, el cual presenta dificultades para captar sarcasmo, y categorizar de forma errónea términos como “banco” o “dinero” bajo la emoción de confianza, a pesar de que el contexto pueda ser de rechazo.

Finalmente el modelo LDA puede que simplifique el análisis ya que asume que cada comentario pertenece a un solo tópico principal. En realidad, las preocupaciones de los ciudadanos son transversales, es decir que se afectan entre ellas, la inflación influye directamente en la capacidad de pagar un alquiler o en el encarecimiento de la compra.

5.4. Líneas de trabajo futuras

Por lo que observamos bastantes líneas de mejora para el futuro, como la integración de datos estructurados. La siguiente fase podría ser cruzar los datos del análisis de sentimiento con indicadores macroeconómicos. Correlacionar el miedo en el tópico 3 con el aumento de los precios del alquiler que daría una validez estadística más veraz al modelo.

Otra mejora sería la inclusión de datos de otras redes sociales donde podamos encontrar un perfil menos técnico para contrastar si la asfixia económica se repite con la misma intensidad y si las temáticas son similares o podemos encontrar diferencias entre las preocupaciones según el perfil.

La meta final sería construir un prototipo de la aplicación que integre estos hallazgos, testeando si realmente todas estas herramientas que hemos mencionado logran efectivamente reducir los niveles de estrés y tristeza detectados en este estudio.

La investigación realizada confirma que la sociedad española, representada en este estudio por los usuarios de comunidades financieras digitales, atraviesa una profunda crisis de seguridad financiera.

El estudio revela que el ciudadano no padece únicamente de una falta de recursos, sino una desconexión sistemática con las herramientas que podrían protegerlos y una desconexión con el estado que no ven como un protector. Mientras la inflación y la vivienda se disparan y llegan a niveles críticos, la gente actúa a ciegas sin ninguna estrategia y el conocimiento necesario para combatirlos permanece restringido a nichos de alta alfabetización.

Ante este escenario, la creación de una solución digital que integre educación financiera interactiva, apoyo fiscal y planificación estratégica no es solo una oportunidad, sino una necesidad. Que puede transformar esa capacidad de supervivencia cotidiana en una estructura de inversión sólida basada en herramientas simples de comprender que han demostrado ser las mejores para combatir este tipo de situaciones. De esta forma conseguimos democratizar estas herramientas y sustituir la incertidumbre por métricas de control real, que ofrece al usuario la posibilidad de transitar hacia una independencia financiera sólida.

REFERENCIAS:

Sobrino, R., Sobrino, R., & Sobrino, R. (2025, 4 diciembre). La mayoría de los españoles creen dominar sus finanzas, pero solo uno de cada tres entiende la inflación. *El País*.

https://elpais.com/economia/2025-12-04/la-mayoria-de-los-espanoles-creen-dominar-sus-finanzas-pero-solo-uno-de-cada-tres-entienden-la-inflacion.html?event_log=go%E2%80%9D

Moreno, M. G. (2025, 6 octubre). Falta educación financiera: el 62% de los jóvenes no invierte porque no sabe cómo hacerlo. *elEconomista.es*.

<https://www.eleconomista.es/actualidad/noticias/13570579/09/25/falta-educacion-financiera-el-62-de-los-jovenes-no-invierte-porque-no-sabe-como-hacerlo.html>

Funcas - Fundación de las Cajas de Ahorros. (2023, 5 diciembre). *La educación financiera de los españoles a examen: Encuesta de Competencias*

Financieras - Funcas. Funcas.

<https://www.funcas.es/odf/la-educacion-financiera-de-los-espanoles-a-examen-encuesta-de-competencias-financieras/>

Fernández, D., Fernández, D., & Fernández, D. (2024, 19 febrero). 'Finfluencers': estos son los peligros de los gurús financieros de los que se fían sus hijos. *El País*.

<https://elpais.com/economia/negocios/2024-02-19/finfluencers-estos-son-los-peligros-de-los-gurus-financieros-a-los-que-siguen-los-miembros-de-la-generacion-z.html>

CarlosAF, & CarlosAF. (2025, 10 diciembre). *Seis de cada diez españoles que no invierten en productos con riesgo lo hacen por falta de educación financiera*. El Asesor Financiero.

<https://elasesorfinanciero.com/seis-de-cada-diez-espanoles-que-no-invierten-en-productos-con-riesgo-lo-hacen-por-falta-de-educacion-financiera/>

ANEXO

El código siguiente permite la reproducción total del estudio. Se incluye una estructura que abarca desde la extracción de datos (scraping) hasta la generación de las métricas de sentimiento y bigramas que sustentan la propuesta de la App.

Los conjuntos de datos limpios, las listas de *stopwords* personalizadas y el script original están disponibles en el siguiente repositorio.

https://disco.uv.es/disco?fileman:ca:principal:user:disco/Un_puente_digital::1at

