



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

CAMPUS D'ALCOI

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Politécnica Superior de Alcoy

Desarrollo de un software de análisis de estado anímico
con perspectiva de su uso en el ámbito empresarial

Trabajo Fin de Grado

Grado en Ingeniería Informática

AUTOR/A: Soler Garcia, Alvaro

Tutor/a: Guerola Navarro, Vicente

CURSO ACADÉMICO: 2022/2023

TÍTULO: Desarrollo de un software de análisis de estado anímico con perspectiva de su uso en el ámbito empresarial

Resumen

Con el paso de los años cada vez dotamos de más importancia al perfil anímico y psicológico de las personas, es un elemento clave en multitud de aspectos y competencias que durante gran parte de la historia de la humanidad se ha pasado por alto o ignorado. Es gracias a nuestras emociones y personalidades por lo que hacemos las cosas que decidimos hacer y reaccionamos de diversas maneras ante las situaciones que se nos plantean en el día; en general, son las que dictan nuestros patrones de comportamiento.

El conocimiento de estos patrones, cómo identificarlos, trabajar con ellos y gestionarlos presenta muchas opciones de uso y posibilidades para multitud de áreas del conocimiento y aplicaciones, por lo que evidentemente ha encontrado también su lugar dentro del mundo empresarial. Las empresas comenzaron a darse cuenta de que un aspecto vital del proceso de compra de un cliente está influenciado por su perfil psicológico, al igual que la productividad y satisfacción de un empleado también lo está, por lo que empezaron a desarrollar estrategias y prácticas que incluyeran este factor.

Por tanto, con este Trabajo de Fin de Grado busco desarrollar una herramienta que de manera sencilla permita extraer conclusiones iniciales acerca del estado anímico de una persona para que sean utilizadas en posteriores ámbitos de uso, tales como estudios de segmentación de público objetivo o gestión del cambio dentro de la organización. Para ello se desarrolla un software de análisis de sentimiento que posee un modelo entrenado en el lenguaje que toma el texto producido por la persona en su perfil de Twitter y lo analiza para mostrar dichas conclusiones.

Acompañando al desarrollo del software, también se estudia la viabilidad de generar negocio con él y que sea rentable mediante las herramientas ofrecidas por Porter así como el estudio de las fortalezas y debilidades del proyecto y la planificación del mismo en cuanto a recursos financieros y temporales.

Palabras clave: Análisis de sentimientos; Python; Redes Sociales; Aprendizaje; Perfil psicológico; Modelo entrenado en lenguaje; Viabilidad económica

TITLE: Development of a state of mind analysis software with a perspective of its use in the business environment

Abstract:

Over the years, we increasingly give more importance to the mental and psychological profile of people, it is a key element in many aspects and competencies that for much of human history has been overlooked or ignored. It is thanks to our emotions and personalities that we do the things that we decide to do and we react in various ways to the situations that arise in our day; in general, they are the ones that dictate our behavior patterns.

The knowledge of these patterns, how to identify them, work with them and manage them presents many options for use and possibilities for a multitude of areas of knowledge and applications, so it has obviously also found its place in the business world. Companies began to realize that a vital aspect of a customer's buying process is influenced by their psychological profile, just as the productivity and satisfaction of an employee is also influenced, so they began to develop strategies and practices that included this factor.

Therefore, with this Final Degree Project I seek to develop a tool that in a simple way allows to draw initial conclusions about a person's state of mind so that they can be used in subsequent areas of use, such as target audience segmentation studies or management of the change within the organization. For this, a software for analysis of

An understanding that has a model trained in the language that takes the text produced by the person on their Twitter profile and analyzes it to show these conclusions.

Accompanying the development of the software, the feasibility of generating business with it and making it profitable through the tools offered by Porter is also studied, as well as the study of the strengths and weaknesses of the project and its planning in terms of financial and temporal resources.

Keywords: Sentiment analysis; Python; Social networks; Learning; Psychological profile; Model trained in language; Economic feasibility

TÍTOL: Desenvolupament d'un programari d'anàlisi d'estat anímic amb perspectiva del seu ús en l'àmbit empresarial

Resum

Amb el pas dels anys cada vegada dotem de més importància al perfil anímic i psicològic de les persones, és un element clau en multitud d'aspectes i competències que durant gran part de la història de la humanitat s'ha passat per alt o ignorat. És gràcies a les nostres emocions i personalitats pel que fem les coses que decidim fer i reaccionem de diverses maneres davant les situacions que se'ns plantegen en el dia; en general, són les que dicten els nostres patrons de comportament.

El coneixement d'aquests patrons, com identificar-los, treballar amb ells i gestionar-los presenta moltes opcions d'ús i possibilitats per a multitud d'àrees del coneixement i aplicacions, per la qual cosa evidentment ha trobat també el seu lloc dins del món empresarial. Les empreses van començar a adonar-se que un aspecte vital del procés de compra d'un client està influenciat pel seu perfil psicològic, igual que la productivitat i satisfacció d'un empleat també ho està, per la qual cosa van començar a desenvolupar estratègies i pràctiques que inclogueren aquest factor.

Per tant, amb aquest Treball de Fi de Grau busque desenvolupar una eina que de manera senzilla permeti extraure conclusions inicials sobre l'estat anímic d'una persona perquè siguin utilitzades en posteriors àmbits d'ús, com ara estudis de segmentació de públic objectiu o gestió del canvi dins de l'organització. Per a això es desenvolupa un programari d'anàlisi de *entimiento que posseeix un model entrenat en el llenguatge que pren el text produït per la persona en el seu perfil de Twitter i l'analitza per a mostrar aquestes conclusions.

Acompanyant al desenvolupament del programari, també s'estudia la viabilitat de generar negoci amb ell i que siga rendible mitjançant les eines oferides per *Porter així com l'estudi de les fortaleses i febleses del projecte i la planificació del mateix quant a recursos financers i temporals.

Paraules clau: Anàlisi de sentiments; Python; Xarxes Socials; Aprenentatge; Perfil psicològic; Model entrenat en llenguatge; Viabilitat econòmic

Contenido

1.INTRODUCCIÓN	7
2.REVISIÓN DE LITERATURA Y FUNDAMENTOS TEÓRICOS	13
3.ANÁLISIS DE NEGOCIO Y DESARROLLO ESTRATÉGICO.....	32
4.PROPOSTA DEL SOFTWARE	50
5.REVISIÓN DE LAS TECNOLOGÍAS Y MEDIOS UTILIZADOS.....	55
6.PLANTEAMIENTO DE LA SOLUCIÓN.....	61
7. PLANIFICACIÓN DEL PROYECTO UTILIZANDO LA METODOLOGÍA SCRUM.....	75
8.POSIBLES MEJORES DE LA IMPLEMENTACIÓN	86
9.ÁMBITO ÉTICO-LEGAL	95
10.IMPLEMENTACIÓN DE LA ESTRATEGIA	99
11.RESULTADOS INICIALES DEL PROYECTO.....	102
12.CONCLUSIONES	104
13.BIBLIOGRAFÍA	106
14.ANEXOS	108

1.INTRODUCCIÓN

1.1 Contexto

Desde el inicio de la humanidad, las personas han desarrollado inherentemente una serie de capacidades de carácter social y necesidades ligadas a ellas, siendo estas una manifestación de su conducta, pensamientos e inquietudes así como un impulso a sus motivaciones y objetivos.

El autor Lozares-Colina desarrolla en su teoría de las redes sociales que esta manera que tienen los seres humanos de interactuar y hacer uso de sus habilidades sociales, se realiza en un espacio o “red” conformado por un conjunto de personas y el entorno en el que estos actores desarrollan sus relaciones.

“Las Redes Sociales pueden definirse como un conjunto bien delimitado de actores, individuos, grupos, organizaciones, comunidades, sociedades globales etc. – vinculados unos a otros a través de una relación o un conjunto de relaciones sociales” (Lozares-Colina, 1996, pág 108).

Con el transcurso de los siglos y el desarrollo cultural y tecnológico de las sociedades, estas “redes sociales” han sido sometidas a continuo cambio y evolución, hasta el punto de que con la llegada de Internet han llegado plataformas desarrolladas única y exclusivamente con el propósito de socializar.

Estas plataformas imitan y potencian las relaciones sociales tradicionales, fueran a distancia o no y ofrecen además añadidos a las mismas, por lo que se han convertido en una manera de relacionarse muy popular entre la población mundial y prácticamente la totalidad de las personas que viven con acceso a Internet y a las redes sociales han acabado por integrarlas en su vida, eligiendo la que mejor se adapte a ellos según sus preferencias.

Las personas cuando se relacionan con otras muestran de muchas maneras aspectos clave acerca de sus vidas: si están inquietos, felices, preocupados, cansados etc. En ámbito general, expresan cómo se sienten en todo momento, y depende de ello el cómo esa persona reaccionará o dirigirá sus acciones del día a día.

Por ello, estas plataformas se han convertido en una fuente inagotable de interacción humana, son bases de datos enormes de sentimientos y comportamientos, que con un análisis adecuado pueden aportar conocimiento muy valioso para múltiples áreas de investigación.

1.2 Introducción al análisis de sentimiento

Una de las herramientas más utilizadas para explotar estas fuentes de información originadas es el análisis de sentimiento o la minería de opinión. Esto no es otra cosa que el procesamiento del lenguaje natural, su análisis y procesamiento lingüístico de manera computacional para obtener e identificar información subjetiva de los recursos analizados mediante la estadística.

De manera general, esto se logra mediante el uso de técnicas que separan el texto en subsecuencias denominadas *tokens* que se analizan y comparan en bases de datos que asignan por estadística cuando una expresión es utilizada para expresar un sentimiento concreto. Una vez se ha realizado la comparación, se asignan una escala de valores a cada subsecuencia y se calcula un balance general de la intencionalidad del mensaje completo.

Es importante destacar que el análisis de sentimiento puede ser más complejo que simplemente asignar una polaridad positiva, negativa o neutra a un texto. Algunos enfoques avanzados pueden considerar la intensidad de los sentimientos expresados, así como la detección de sentimientos mixtos o ambivalentes. Además, el análisis de sentimiento puede adaptarse a contextos específicos, como la detección de opiniones sobre productos, servicios o eventos particulares.

Además, hay que tener en cuenta que el análisis de sentimiento basado en el procesamiento del lenguaje natural y la estadística puede tener limitaciones. Puede haber ambigüedad en el lenguaje, sarcasmo, ironía o expresiones idiomáticas que pueden afectar la precisión de los resultados. Además, los sentimientos pueden variar según el contexto cultural y lingüístico, lo que requiere adaptaciones específicas.

En resumen, el análisis de sentimiento es una técnica que permite extraer información subjetiva y emocional de los textos utilizando herramientas computacionales. A través de la comparación y valoración de los tokens en bases de datos de sentimientos, se obtiene una medida global del tono emocional expresado en el texto analizado. Esta técnica es utilizada en una variedad de aplicaciones, desde la gestión de la reputación en línea hasta la investigación de mercados y la toma de decisiones empresariales.

1.3 Objetivos

El objetivo de este TFG es el desarrollo de un proyecto de software primigenio de análisis de sentimiento que pueda interesar a diferentes empresas para invertir y desarrollar un producto final adaptada a sus necesidades y objetivos.

Para ello se desarrollarán diversos análisis de la viabilidad del proyecto así como sus puntos fuertes y débiles dentro del mercado y se expondrán tecnologías utilizadas y desarrollo del propio software, así como posibles usos finales y adaptaciones a necesidades concretas.

1.3.1 Objetivos a corto plazo

El principal objetivo a corto plazo es el desarrollo de un software inicial capaz de llevar a cabo el procesamiento de texto para su posterior análisis del sentimiento y obtención de resultados iniciales con los que ofrecer información convertible en conocimiento para las empresas.

1.3.2 Objetivos a largo plazo

En cuanto a los objetivos a largo plazo, estos se definen bajo el supuesto de ser finalmente proveedor de este servicio de una empresa. En ese momento, el software llevará su desarrollo a un siguiente nivel: desarrollo de interfaces personalizadas a las necesidades de la empresa, elección de la fuente de información objetivo de las estrategias de la empresa, alineación con sus objetivos y en general, satisfacer todas las necesidades planteadas por el cliente para su desarrollo empresarial.

1.4 Alcance del trabajo

Una vez definidos los objetivos de este TFG, cabe redactar la estructura que se va a seguir para la elaboración de este trabajo y proyecto.

De manera previa al inicio del desarrollo, se va a añadir un apartado enfocado a la revisión de literatura, con el que incluir lecturas relacionadas con el tema tratado y que sirven de fuente de información y contexto para formar un punto de partida más sólido, sobre todo porque el análisis de sentimientos es un área de conocimiento e investigación de gran envergadura que contrae una complejidad notoria si se desea entrar en ella en profundidad.

En el momento que se ha formado ese punto de inicio contextualizado, da comienzo el estudio de la estrategia a seguir para el desarrollo del proyecto y su viabilidad.

En primer lugar se procede a un análisis del entorno con el objetivo de encontrar las debilidades y amenazas presentes para el proyecto así como visualizar las oportunidades y fortalezas que va a poder explotar el mismo para lograr los objetivos planteados.

En segundo lugar y de antecedente a la elección de la estrategia a seguir, se definirá tanto la misión como la visión y los objetivos estratégicos que el proyecto debe seguir y quienes

son los principales grupos de interés existentes en un proyecto como el nuestro y cómo podemos atraerlos.

A continuación, procederá analizar a la competencia y en base a ello elegir la estrategia que mejor se adapte a la misión, visión y objetivos establecidos para conseguir un mayor número de ventajas competitivas que ofrecer en la búsqueda de captar la atención del mayor número posible de posibles clientes para nuestro servicio.

Para finalizar, una vez tomada la decisión de qué estrategia seguir, comenzará el desarrollo del proyecto y la implementación de la estrategia en el mismo. Además, para el desarrollo del mismo se estudiarán las tecnologías más comunes en el área y se tomará como base y fuente de texto la red social Twitter para el desarrollo del software, pues entre todas las existentes, dado su formato de plataforma, es de las que más texto alberga y está diseñada para que sus usuarios la utilicen para expresar lo que piensan en cada momento mediante texto, dando así multitud de opiniones, recuerdos e interacciones humanas analizables por nuestro proyecto.

1.5 Motivación

Mi interés por la psicología humana ha estado siempre presente, hasta el punto que era una de las opciones que tanteaba a la hora de elegir qué carrera estudiar. Es un tema que siempre me ha apasionado y despertado mucho interés: cómo reaccionan las personas, cómo sienten en cada circunstancia...

Esto ha ocasionado que me llamen especial atención todas las tecnologías referidas a la psicología, todos los avances y maneras que tiene la tecnología de interactuar con el ámbito. Al fin y al cabo, las posibilidades de la informática rozan lo infinito. De entre todas ellas, me interesó en especial el análisis de sentimiento, pues según mi apreciación de la sociedad y cómo estamos evolucionando, cada vez somos más conscientes de la importancia de expresar nuestros sentimientos y de cuidar nuestra salud mental y a la vez tenemos más dificultad para expresarlos.

Con el crecimiento abismal de las redes sociales, el uso de las pantallas nos ha unido y a la vez distanciado del resto de seres humanos, de tal forma que acabamos que acabamos teniendo dificultades para mostrar nuestros sentimientos y las redes acaban llenas de personas pidiendo ayuda a gritos con la boca cerrada. Por eso considero que una mayor conciencia de estos aspectos y un uso responsable por parte de las empresas de tecnologías como el análisis de sentimiento podrían contribuir mucho a la salud mental global y prevenir muchos problemas.

Sin embargo, claro está, no es ese su único uso, también pueden usarse con fines más “optimistas” o no tan dramáticos, y muchos de ellos con finalidad económica, por lo que me parece un punto interesante que explorar y que además relaciona las dos carreras que he elegido estudiar a la vez, la mezcla del mundo de la informática y la economía y la empresa en un tema que me interesa.

2.REVISIÓN DE LITERATURA Y FUNDAMENTOS TEÓRICOS

Durante este apartado se recogerá la búsqueda previa al desarrollo del proyecto de artículos, documentos y fundamento teóricos relacionados con el análisis del sentimiento al que está enfocado este Trabajo de Final de Grado. Estos documentos e información servirán como base de entendimiento del tema tratado.

2.1 Análisis de sentimiento: conceptos y definiciones

El análisis de sentimiento es una disciplina que se centra en el procesamiento del lenguaje natural y tiene como objetivo identificar y comprender las emociones, opiniones y actitudes expresadas en el texto. Consiste en determinar si un texto o fragmento de texto transmite un sentimiento positivo, negativo o neutro, y en algunos casos, también evaluar la intensidad de dicho sentimiento.

En el análisis de sentimiento, se utiliza una combinación de técnicas lingüísticas, estadísticas y de aprendizaje automático para extraer información subjetiva y emocional de los textos. Una de las técnicas más comunes es el enfoque basado en reglas, que consiste en utilizar un conjunto de reglas predefinidas para asignar polaridades a palabras y frases y calcular una polaridad global para el texto analizado.

Además del enfoque basado en reglas, existen otros enfoques más avanzados en el análisis de sentimiento, como la clasificación supervisada y el aprendizaje automático. La clasificación supervisada implica entrenar un modelo utilizando un conjunto de datos etiquetados previamente con sentimientos conocidos, y luego utilizar ese modelo para clasificar nuevos textos. Por otro lado, el aprendizaje automático permite que un modelo aprenda de forma automática a partir de un conjunto de datos sin necesidad de una supervisión explícita.

Es importante destacar que el análisis de sentimiento no se limita a la asignación de una polaridad positiva, negativa o neutra. También se puede considerar la intensidad de los sentimientos expresados, lo que permite capturar matices y variaciones en la emocionalidad de un texto. Además, el análisis de sentimiento puede abordar la detección de sentimientos mixtos, donde un texto puede contener elementos tanto positivos como negativos.

El análisis de sentimiento tiene aplicaciones en diversos campos, como el análisis de la opinión de los usuarios en redes sociales, la monitorización de la reputación en línea de una marca o empresa, la investigación de mercados, el análisis de campañas políticas y la detección de tendencias y eventos en tiempo real. El análisis de sentimiento en Twitter, en particular, ha ganado relevancia debido a la gran cantidad de usuarios y la naturaleza pública y rápida de los mensajes en esta plataforma.

En conclusión, el análisis de sentimiento es una disciplina que se enfoca en la identificación y comprensión de las emociones y opiniones expresadas en el texto. Se utilizan técnicas lingüísticas, estadísticas y de aprendizaje automático para extraer información subjetiva y emocional de los textos, y se pueden considerar aspectos como la polaridad, la intensidad y la detección de sentimientos mixtos. El análisis de sentimiento tiene diversas aplicaciones en diferentes áreas y el análisis de sentimiento en Twitter ha adquirido importancia debido a su popularidad y dinamismo.

2.2 Aplicaciones del análisis de sentimiento en redes sociales

Las redes sociales, como Twitter, han experimentado un crecimiento exponencial en los últimos años, convirtiéndose en una fuente inagotable de información generada por los usuarios. El análisis de sentimiento ha encontrado un hueco en este contexto, ya que permite comprender las opiniones y emociones de los usuarios expresadas en sus publicaciones. A continuación, se explorarán algunas de las aplicaciones más relevantes del análisis de sentimiento en redes sociales, con un enfoque particular en Twitter.

1. **Toma de decisiones empresariales:** El análisis de sentimiento en Twitter brinda a las empresas una visión valiosa sobre cómo los usuarios perciben sus productos, servicios o campañas de marketing. Al analizar el sentimiento de los tweets relacionados con su marca, las empresas pueden evaluar la satisfacción del cliente, identificar problemas y áreas de mejora, y tomar decisiones informadas para adaptar sus estrategias y mejorar la experiencia del cliente.
2. **Gestión de la reputación y repercusión:** El análisis de sentimiento en redes sociales, incluyendo Twitter, permite a las empresas monitorear y gestionar su marca online. Pueden detectar y responder rápidamente a comentarios negativos o quejas de los usuarios, evitando posibles escándalos. Asimismo, el análisis de sentimiento proporciona información sobre la percepción de la marca y el impacto de las acciones de relaciones públicas y marketing en la comunidad online.
3. **Investigación de mercados:** El análisis de sentimiento en Twitter se utiliza en la investigación de mercados para comprender las preferencias y tendencias de los consumidores. Al analizar los sentimientos expresados en los tweets relacionados con productos o servicios específicos, se pueden identificar patrones sobre las percepciones de los consumidores, sus necesidades y expectativas. Esto ayuda a las empresas a desarrollar estrategias de mercado más efectivas y a tomar decisiones basadas en datos.

4. **Análisis de campañas políticas:** El análisis de sentimiento en Twitter también se aplica al ámbito político. Permite analizar el sentimiento de los usuarios hacia los candidatos, partidos políticos y temas relevantes. Esta información es especialmente útil durante las campañas electorales, ya que permite a los políticos y sus equipos comprender la opinión pública, identificar los problemas más importantes para los votantes y ajustar su estrategia de comunicación en consecuencia.

5. **Detección de tendencias y eventos en tiempo real:** Twitter es conocido por su naturaleza en tiempo real, lo que lo convierte en una plataforma valiosa para detectar tendencias y eventos emergentes. El análisis de sentimiento en Twitter permite identificar rápidamente los temas más populares y las reacciones emocionales asociadas a ellos. Esto es útil para periodistas, empresas de medios de comunicación y agencias de inteligencia que desean estar al tanto de los acontecimientos actuales y las conversaciones en línea.

En resumen, el análisis de sentimiento en Twitter ofrece diversas aplicaciones prácticas en distintos ámbitos. Permite a las empresas evaluar la percepción de sus productos y servicios, gestionar su reputación en línea y tomar decisiones informadas. También es útil en la investigación de mercados, el análisis de campañas políticas y la detección de tendencias en tiempo real. En conjunto, el análisis de sentimiento en Twitter proporciona una visión invaluable de las opiniones y emociones de los usuarios, brindando información relevante para la toma de decisiones y la comprensión de la comunidad en línea.

2.3 Estado del arte en el análisis de sentimiento en Twitter

El análisis de sentimiento en Twitter ha experimentado avances significativos en los últimos años, gracias al desarrollo de diversas técnicas y enfoques. A continuación, se presentan algunos de los avances más relevantes del análisis de sentimiento en Twitter:

Aprendizaje profundo: Los modelos de aprendizaje profundo, como las redes neuronales recurrentes (RNN) y las redes neuronales convolucionales (CNN), han demostrado ser eficaces en el análisis de sentimiento en Twitter. Estos modelos pueden capturar mejor las relaciones contextuales y las características semánticas de los textos, lo que mejora la precisión del análisis de sentimiento. Algunos autores y artículos que hablan sobre ellos:

Tang, D., Qin, B., Liu, T., & Yang, Y. (2015). Document modeling with gated recurrent neural network for sentiment classification

Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning word vectors for sentiment analysis

Transferencia de aprendizaje: La transferencia de aprendizaje es otra técnica utilizada en el análisis de sentimiento en Twitter. Consiste en aprovechar los conocimientos adquiridos en una tarea relacionada, como el análisis de sentimiento en texto general, y aplicarlos al análisis de sentimiento en Twitter. Esto permite obtener mejores resultados con conjuntos de datos más pequeños y reduce la necesidad de etiquetar grandes volúmenes de datos específicos de Twitter. Algunos autores y artículos que hablan sobre ellos:

Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification

Modelos de lenguaje pre entrenados: Los modelos de lenguaje pre entrenados, como BERT (Bidirectional Encoder Representations from Transformers) y GPT (Generative Pre-trained Transformer), han demostrado ser altamente efectivos en el análisis de sentimiento en Twitter. Estos modelos capturan las relaciones contextuales y la semántica

de las palabras y frases, lo que mejora la comprensión del sentimiento en los tweets. Algunos autores y artículos que hablan sobre ellos:

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners

Enfoques basados en aspectos: Los enfoques basados en aspectos se centran en identificar el sentimiento en relación con aspectos específicos, como características de productos o servicios. Estos enfoques utilizan técnicas de extracción de aspectos y análisis de sentimiento en cada aspecto individualmente, lo que proporciona una visión más detallada y precisa de las opiniones de los usuarios. Algunos autores y artículos que hablan sobre ellos:

Quintero, Y. C., & García, L. A. ESTUDIO DEL ANÁLISIS DE SENTIMIENTOS BASADO EN ASPECTOS.

Análisis de sentimiento multimodal: El análisis de sentimiento multimodal combina información textual con otros tipos de datos, como imágenes o videos, para mejorar la comprensión del sentimiento en Twitter. Por ejemplo, se pueden utilizar técnicas de análisis de emociones en imágenes o análisis de expresiones faciales para complementar el análisis de sentimiento basado en texto. Algunos autores y artículos que hablan sobre ellos:

Cabada, R. Z., Estrada, M. L. B., & López, H. M. C. (2019). Reconocimiento multimodal de emociones orientadas al aprendizaje. Research in Computing Science, 148(7), 153-165.

Métodos de etiquetado supervisado: Dado que etiquetar grandes conjuntos de datos puede ser costoso y consumir mucho tiempo, se han propuesto métodos de etiquetado supervisado en el análisis de sentimiento en Twitter. Estos métodos aprovechan tanto datos etiquetados como no etiquetados para entrenar modelos de manera más eficiente y mejorar el rendimiento del análisis de sentimiento. Algunos autores y artículos que hablan sobre ellos:

Rodríguez Tapia, S., & Camacho-Cañamón, J. (2018). Los métodos de aprendizaje automático supervisado en la clasificación textual según el grado de especialización.

Morales Rodríguez, J. (2021). Evaluación de métodos auto-supervisados y semi-supervisados para la extracción de características visuales en el contexto de recuperación de imágenes basada en dibujos.

A partir de estos modelos, métodos y enfoques, sumado a los artículos citados, se procede a expandir cada uno de ellos en los siguientes apartados de manera que se presente una comparativa esquemática similar entre ellos para ver algunos de sus puntos fuertes y débiles.

2.3.1 Aprendizaje profundo en el análisis de sentimiento en Twitter

En el campo del análisis de sentimiento en Twitter, el aprendizaje profundo ha surgido como una técnica poderosa y efectiva. Los modelos de aprendizaje profundo, como las redes neuronales recurrentes (RNN) y las redes neuronales convolucionales (CNN), han demostrado su capacidad para capturar patrones complejos en los textos de los tweets y mejorar la precisión del análisis de sentimiento.

Las RNN son especialmente adecuadas para el análisis de secuencias de texto, ya que pueden modelar dependencias a largo plazo entre las palabras en un tweet. Estas redes están diseñadas para aprender y recordar información contextual de los tweets anteriores mientras procesan uno nuevo. Esta memoria contextual permite comprender mejor el sentimiento expresado en un tweet, ya que se pueden capturar las sutilezas y los cambios en el tono a medida que se avanza en la secuencia de palabras.

Por otro lado, las CNN se destacan en la extracción de características relevantes al aplicar filtros convolucionales sobre el texto del tweet. Estos filtros son capaces de detectar patrones locales en los datos, como n-gramas y combinaciones de palabras, lo que permite capturar características semánticas importantes relacionadas con el sentimiento. Las CNN

también son eficientes computacionalmente, lo que las convierte en una opción atractiva para el análisis de sentimiento en grandes volúmenes de tweets.

Además de las RNN y las CNN, también se han desarrollado arquitecturas más complejas que combinan ambos enfoques, como las redes neuronales convolucionales recurrentes (CRNN) y las redes neuronales transformer. Estos modelos aprovechan las fortalezas de cada enfoque para capturar tanto la información contextual como las características semánticas de los tweets.

El aprendizaje profundo ha demostrado su eficacia en el análisis de sentimiento en Twitter al superar los enfoques tradicionales basados en características manuales y reglas predefinidas. La capacidad de estos modelos para aprender automáticamente representaciones de alto nivel a partir de los datos ha llevado a mejoras significativas en la precisión del análisis de sentimiento. Además, el aprendizaje profundo es altamente flexible y puede adaptarse a diferentes dominios y idiomas sin necesidad de ajustes manuales extensivos.

Sin embargo, el aprendizaje profundo también presenta desafíos. Estos modelos requieren grandes conjuntos de datos etiquetados para un entrenamiento efectivo, lo que puede ser costoso y consume tiempo. Además, la interpretación de los resultados de los modelos de aprendizaje profundo puede ser más compleja en comparación con los enfoques tradicionales, lo que puede dificultar la comprensión de los factores que contribuyen al sentimiento en un tweet específico.

A pesar de estos desafíos, el aprendizaje profundo sigue siendo una herramienta prometedora en el análisis de sentimiento en Twitter y continúa siendo objeto de investigación y desarrollo activos. La aplicación de técnicas de aprendizaje profundo en combinación con conjuntos de datos adecuados y una adecuada configuración de parámetros puede proporcionar resultados más precisos y enriquecer nuestra comprensión del sentimiento expresado en los tweets.

2.3.2 Transferencia de aprendizaje en el análisis de sentimiento en Twitter

La transferencia de aprendizaje es una técnica ampliamente utilizada en el análisis de sentimiento en Twitter. Consiste en aprovechar los conocimientos adquiridos en una tarea relacionada, como el análisis de sentimiento en texto general, y aplicarlos al análisis de sentimiento en Twitter. Esto permite obtener mejores resultados con conjuntos de datos más pequeños y reduce la necesidad de etiquetar grandes volúmenes de datos específicos de Twitter.

En el contexto del análisis de sentimiento en Twitter, la transferencia de aprendizaje se utiliza de varias formas. Una de ellas es utilizar modelos pre entrenados en conjuntos de datos generales, como noticias o reseñas de películas, y luego adaptarlos a los datos de Twitter. Estos modelos pre entrenados capturan características y patrones lingüísticos generales que también pueden ser relevantes en el análisis de sentimiento en Twitter. Al adaptarlos a los datos específicos de Twitter, los modelos pueden aprender a reconocer las particularidades y las expresiones de sentimiento propias de esta plataforma.

Otra forma de aplicar la transferencia de aprendizaje en el análisis de sentimiento en Twitter es utilizar conjuntos de datos etiquetados en tareas similares, como el análisis de sentimiento en otras redes sociales o plataformas. Los conocimientos adquiridos en estas tareas pueden transferirse al análisis de sentimiento en Twitter, ya que comparten características y desafíos comunes. Esto permite beneficiarse de los conjuntos de datos más grandes y variados disponibles en otras plataformas, sin la necesidad de etiquetar grandes volúmenes de datos específicos de Twitter.

La transferencia de aprendizaje también puede aplicarse a nivel de características. En lugar de utilizar directamente los textos de los tweets, se pueden extraer características generales, como la presencia de ciertas palabras o la frecuencia de emoticonos, y luego utilizar estas características en el análisis de sentimiento en Twitter. Estas características pueden ser aprendidas de conjuntos de datos generales y aplicarse posteriormente a los datos de Twitter.

Una de las ventajas clave de la transferencia de aprendizaje en el análisis de sentimiento en Twitter es que reduce la dependencia de grandes conjuntos de datos específicos de esta plataforma. Al aprovechar los conocimientos adquiridos en tareas relacionadas, se pueden obtener resultados prometedores incluso con conjuntos de datos más pequeños. Esto es especialmente útil en escenarios en los que la disponibilidad de datos etiquetados en Twitter es limitada o costosa.

No obstante, es fundamental considerar que la transferencia de aprendizaje puede no ser viable en todas las situaciones. Esta técnica está condicionada a la similitud entre las tareas relacionadas y la disponibilidad de conjuntos de datos adecuados. Además, la adaptación de modelos pre entrenados a los datos de Twitter puede requerir ajustes y refinamientos adicionales para capturar las particularidades propias de esta plataforma.

2.3.2 Transferencia de aprendizaje en el análisis de sentimiento en Twitter

La transferencia de aprendizaje consiste en aprovechar los conocimientos adquiridos en una tarea relacionada, como el análisis de sentimiento en texto general, y aplicarlos al análisis de sentimiento en Twitter. Esto permite obtener mejores resultados con conjuntos de datos más pequeños y reduce la necesidad de etiquetar grandes volúmenes de datos específicos de Twitter.

En el contexto del análisis de sentimiento en Twitter, la transferencia de aprendizaje se utiliza de varias formas. Una de ellas es utilizar modelos pre entrenados en conjuntos de datos generales, como noticias o reseñas de películas, y luego adaptarlos a los datos de Twitter. Estos modelos pre entrenados capturan características y patrones lingüísticos generales que también pueden ser relevantes en el análisis de sentimiento en Twitter. Al adaptarlos a los datos específicos de Twitter, los modelos pueden aprender a reconocer las particularidades y las expresiones de sentimiento propias de esta plataforma.

Otra forma de aplicar la transferencia de aprendizaje en el análisis de sentimiento en Twitter es utilizar conjuntos de datos etiquetados en tareas similares, como el análisis de

sentimiento en otras redes sociales o plataformas. Los conocimientos adquiridos en estas tareas pueden transferirse al análisis de sentimiento en Twitter, ya que comparten características y desafíos comunes. Esto permite beneficiarse de los conjuntos de datos más grandes y variados disponibles en otras plataformas, sin la necesidad de etiquetar grandes volúmenes de datos específicos de Twitter.

La transferencia de aprendizaje también puede aplicarse a nivel de características. En lugar de utilizar directamente los textos de los tweets, se pueden extraer características generales, como la presencia de ciertas palabras o la frecuencia de emoticonos, y luego utilizar estas características en el análisis de sentimiento en Twitter. Estas características pueden ser aprendidas de conjuntos de datos generales y aplicarse posteriormente a los datos de Twitter.

Una de las ventajas clave de la transferencia de aprendizaje en el análisis de sentimiento en Twitter es que reduce la dependencia de grandes conjuntos de datos específicos de esta plataforma. Al aprovechar los conocimientos adquiridos en tareas relacionadas, se pueden obtener resultados prometedores incluso con conjuntos de datos más pequeños. Esto es especialmente útil en escenarios en los que la disponibilidad de datos etiquetados en Twitter es limitada.

2.3.3 Modelos de lenguaje pre entrenados en el análisis de sentimiento en Twitter

Los modelos de lenguaje pre entrenados han demostrado ser altamente efectivos en el análisis de sentimiento en Twitter. Estos modelos, como BERT (Bidirectional Encoder Representations from Transformers) y GPT (Generative Pre-trained Transformer), han revolucionado el campo del procesamiento del lenguaje natural al capturar las relaciones contextuales y la semántica de las palabras y frases.

En el análisis de sentimiento en Twitter, los modelos de lenguaje pre entrenados se utilizan para comprender el contexto y la intención detrás de los tweets. Estos modelos se entrenan en grandes cantidades de texto proveniente de diversas fuentes, lo que les permite capturar patrones lingüísticos complejos y entender mejor el significado de las palabras y oraciones en diferentes contextos.

Al aplicar estos modelos de lenguaje pre entrenados en el análisis de sentimiento en Twitter, se obtiene una ventaja significativa. Estos modelos pueden comprender el sentimiento expresado en los tweets al capturar las sutilezas y matices del lenguaje. Además, son capaces de reconocer la ironía, el sarcasmo y otras formas de expresión que son comunes en los mensajes de Twitter.

El uso de modelos de lenguaje pre entrenados en el análisis de sentimiento en Twitter también ofrece la ventaja de no tener que entrenar modelos desde cero. Estos modelos ya han aprendido una representación general del lenguaje y pueden ser finamente ajustados para tareas específicas, como el análisis de sentimiento en Twitter. Esto reduce el tiempo y los recursos necesarios para desarrollar modelos efectivos y acelera el proceso de implementación.

Además, los modelos de lenguaje pre entrenados se pueden combinar con técnicas de transferencia de aprendizaje para adaptarlos específicamente al análisis de sentimiento en Twitter. Esto implica ajustar los modelos pre entrenados en conjuntos de datos

etiquetados de Twitter para capturar las características y las expresiones de sentimiento propias de esta plataforma.

Es importante destacar que los modelos de lenguaje pre entrenados también plantean consideraciones importantes en el análisis de sentimiento en Twitter. Aunque son altamente eficaces, su entrenamiento puede estar sesgado hacia conjuntos de datos específicos, lo que puede influir en sus predicciones en situaciones particulares, como en el contexto de Twitter. Por lo tanto, es fundamental tener en cuenta este sesgo y llevar a cabo ajustes adicionales para adaptar los modelos a las características propias de esta plataforma.

2.3.4 Enfoques basados en aspectos en el análisis de sentimiento en Twitter

Los enfoques basados en aspectos en el análisis de sentimiento en Twitter se centran en identificar el sentimiento en relación con aspectos específicos, como características de productos, servicios o eventos. Estos enfoques permiten obtener una visión más detallada y precisa de las opiniones de los usuarios al analizar el sentimiento asociado a aspectos particulares.

Para aplicar enfoques basados en aspectos en el análisis de sentimiento en Twitter, es necesario realizar dos tareas principales: la extracción de aspectos y el análisis de sentimiento en cada aspecto individualmente.

La extracción de aspectos se refiere a la identificación y clasificación de los aspectos relevantes en los tweets. Esto implica identificar las características específicas que los usuarios mencionan en sus publicaciones, como la calidad del servicio, el precio, la facilidad de uso o la atención al cliente. Esta tarea puede realizarse mediante técnicas de procesamiento de lenguaje natural, como la identificación de palabras clave, análisis gramatical o algoritmos de aprendizaje automático.

Una vez que se han extraído los aspectos relevantes, se lleva a cabo el análisis de sentimiento en cada aspecto individualmente. Esto implica aplicar técnicas de análisis de sentimiento, como la clasificación supervisada, para determinar si los tweets expresan un sentimiento positivo, negativo o neutro hacia cada aspecto específico. Esto permite comprender cómo los usuarios perciben cada característica y qué opiniones tienen al respecto.

Los enfoques basados en aspectos en el análisis de sentimiento en Twitter brindan una visión más completa y detallada de las opiniones de los usuarios, ya que no se limitan a una clasificación general del sentimiento en un tweet, sino que analizan el sentimiento en relación con aspectos específicos. Esto es especialmente útil en escenarios donde se desea comprender cómo los usuarios perciben y evalúan características particulares de productos, servicios o eventos.

Algunos puntos negativos residen en la extracción de aspectos precisa puede ser un desafío debido a la variabilidad y la diversidad del lenguaje utilizado en los tweets. Además, la clasificación del sentimiento en cada aspecto puede requerir un etiquetado manual de datos para entrenar modelos específicos. Esto puede ser costoso y consume tiempo, especialmente cuando se considera la cantidad masiva de datos generados en Twitter.

2.3.5 Análisis de sentimiento multimodal en Twitter

El análisis de sentimiento multimodal en Twitter es un enfoque que combina información textual con otros tipos de datos, como imágenes o videos, para mejorar la comprensión del sentimiento expresado en los tweets. Este enfoque reconoce que el contenido textual por sí solo puede ser limitado y que la información visual o auditiva puede proporcionar un contexto adicional para interpretar y comprender el sentimiento de los usuarios.

En el análisis de sentimiento multimodal, se utilizan técnicas de procesamiento de imágenes, procesamiento de señales de audio u otras modalidades para capturar información emocional o contextual de los tweets. Estas técnicas pueden incluir análisis de emociones en imágenes, detección de expresiones faciales, análisis de tono de voz o incluso análisis de gestos en videos.

Por ejemplo, al considerar imágenes, se pueden extraer características visuales relevantes, como el color, la composición o la presencia de elementos específicos que puedan influir en el sentimiento expresado en el texto del tweet. Del mismo modo, al analizar el audio, se pueden identificar características acústicas, como el tono de voz, la entonación o la velocidad de habla, que pueden proporcionar indicios adicionales sobre el sentimiento del usuario.

Al combinar la información textual con datos multimodales, se busca mejorar la precisión y la comprensión del sentimiento en Twitter. Por ejemplo, un tweet que contiene un texto neutral podría expresar sarcasmo o ironía que solo se puede capturar al considerar el contexto visual o auditivo adjunto al tweet.

2.3.6 Métodos de etiquetado supervisado en el análisis de sentimiento en Twitter

En el análisis de sentimiento en Twitter, el etiquetado de grandes conjuntos de datos puede ser costoso y requerir mucho tiempo debido a la necesidad de asignar manualmente etiquetas a cada tweet. Para abordar este desafío, se han propuesto métodos de etiquetado supervisado que aprovechan tanto datos etiquetados como no etiquetados para entrenar modelos de manera más eficiente y mejorar el rendimiento del análisis de sentimiento.

Estos métodos de etiquetado supervisado se basan en la idea de que no todos los tweets en un conjunto de datos necesitan ser etiquetados manualmente. En su lugar, se pueden utilizar técnicas de aprendizaje automático, como algoritmos de clasificación semi-supervisada o de aprendizaje activo, para identificar y etiquetar automáticamente una parte del conjunto de datos no etiquetados.

Uno de los enfoques comunes en el etiquetado supervisado es la clasificación semi-supervisada. Este enfoque combina un conjunto pequeño de datos etiquetados con un conjunto mucho más grande de datos no etiquetados. A partir de los datos etiquetados, se entrena un modelo inicial que se utiliza para asignar etiquetas a los datos no etiquetados. Luego, el modelo se actualiza y se mejora utilizando tanto los datos etiquetados como los datos recién etiquetados. Este ciclo de entrenamiento y etiquetado se repite varias veces para mejorar gradualmente la precisión del modelo.

Otro enfoque es el aprendizaje activo, que implica seleccionar de manera inteligente los tweets más informativos o inciertos para ser etiquetados manualmente. En lugar de etiquetar aleatoriamente tweets no etiquetados, los algoritmos de aprendizaje activo utilizan medidas de incertidumbre o confianza para determinar qué instancias son más beneficiosas para el modelo. Al etiquetar solo los tweets más informativos o inciertos, se maximiza el uso eficiente del tiempo y los recursos humanos dedicados al etiquetado.

Estos métodos de etiquetado supervisado permiten aprovechar los beneficios de los datos no etiquetados y reducir la carga de trabajo asociada con el etiquetado manual. Al utilizar conjuntos de datos más grandes y diversos, se mejora la capacidad del modelo para capturar la variabilidad y los matices del sentimiento en Twitter.

2.4 Desafíos y limitaciones del análisis de sentimiento en Twitter

A modo de resumen, este apartado presenta la recopilación de los principales problemas vistos en cada uno de los enfoques y caminos del análisis de sentimiento en Twitter vistos en apartados anteriores. En primer lugar, el lenguaje utilizado en Twitter es coloquial, con abreviaciones, jerga y errores gramaticales, lo que dificulta la interpretación precisa de los mensajes y afecta la precisión del análisis de sentimiento. Esta característica del lenguaje puede dificultar la tarea de asignar correctamente el sentimiento expresado por los usuarios.

Además, los tweets a menudo contienen ambigüedades y expresiones sarcásticas que pueden llevar a una interpretación errónea del sentimiento. Detectar el sarcasmo es especialmente desafiante y puede afectar la precisión del análisis de sentimiento. La naturaleza breve de los mensajes en Twitter también limita la cantidad de información disponible para comprender el contexto completo, lo que dificulta una interpretación adecuada del sentimiento expresado.

La variabilidad cultural y lingüística de los usuarios de Twitter también representa un desafío. Los diferentes contextos culturales y lingüísticos influyen en la forma en que se expresa el sentimiento, lo que requiere adaptaciones específicas para comprender las particularidades y matices de diferentes comunidades de usuarios.

El volumen y la velocidad de los datos en Twitter representan otro problema para el análisis de sentimiento. La gran cantidad de tweets generados en tiempo real requiere técnicas eficientes para procesar y analizar los datos de manera precisa y oportuna. Además, obtener conjuntos de datos etiquetados para el entrenamiento de modelos de análisis de sentimiento puede ser costoso y laborioso. La disponibilidad de datos de alta calidad con etiquetas precisas es limitada.

Por último, los cambios en el lenguaje y las tendencias en Twitter son frecuentes, lo que implica que los modelos de análisis de sentimiento deben actualizarse y adaptarse

constantemente para mantener su precisión. Es necesario estar al tanto de las nuevas expresiones y evoluciones en el lenguaje utilizado en Twitter para garantizar resultados confiables.

Estas limitaciones subrayan la complejidad del análisis de sentimiento en Twitter. A pesar de ello, sigue siendo una herramienta valiosa para comprender las opiniones y emociones de los usuarios en tiempo real, y se continúa trabajando en el desarrollo de técnicas y enfoques más sofisticados para abordar estas limitaciones y mejorar la precisión del análisis de sentimiento en esta plataforma.

A modo de resumen final, el análisis de sentimiento en Twitter se basa en el procesamiento del lenguaje natural para identificar y extraer información emocional de los textos generados por los usuarios. Aunque tiene aplicaciones en áreas como la toma de decisiones empresariales y la gestión de la reputación, presenta desafíos como el lenguaje coloquial y la variabilidad cultural. A pesar de estas limitaciones, sigue siendo una herramienta valiosa y se busca mejorar su precisión. En el próximo apartado, se explorará el análisis de negocio y desarrollo estratégico del proyecto.

3. ANÁLISIS DE NEGOCIO Y DESARROLLO ESTRATÉGICO

Como paso previo a definir la estrategia que se va a emplear para llevar a cabo este proyecto de cara a comercializar el servicio, conviene realizar un análisis estratégico.

Para ello, se procede a desarrollar dos tipos del mismo, uno para el ámbito externo y otro para el ámbito interno del proyecto. Con ello se busca obtener tanto las debilidades y amenazas como fortalezas y oportunidades que el proyecto tiene y debe de tener en cuenta para un correcto desarrollo.

3.1 Análisis del entorno o análisis externo

El primer paso es desarrollar un análisis del sector que identifique amenazas y oportunidades del proyecto para poder actuar en consecuencia a las mismas.

Este tipo de análisis se caracteriza por centrarse solamente en la parte del entorno que influye de manera directa o indirecta con el proyecto.

Dentro de dicho análisis se distingue entre dos niveles a ser analizados, conocidos comúnmente como macroentorno y microentorno.

3.1.1 Macroentorno

El objetivo de este análisis reside en obtener aquellos aspectos del sistema que influyen en gran medida e impacto con los intereses del proyecto para analizar el grado de repercusión de los mismos.

De entre todas las técnicas que se utilizan para desarrollar un análisis de este tipo vamos a destacar el conocido como análisis PEST .

A. Análisis PEST

Dicho estudio se realiza a nivel nacional, puesto que uno de los objetivos del proyecto es darse a conocer e iniciarse en primer lugar, dentro del ámbito español.

Con todo esto explicado, se procede a analizar las dimensiones mencionadas:

-Político-legales: En primer lugar, destacar que nos encontramos en una situación de crisis política de manera ya extendida por varios años. Cuando todavía no había acabado la crisis pandémica del COVID-19 que dio comienzo el 11 de marzo de 2020 y trajo recesión económica, paro de muchas actividades laborales y su consecuente repercusión del mercado, nos encontramos ahora ante una crisis económica generada por el reciente inicio de una guerra por invasión al este de Europa entre Rusia y Ucrania.

Como era de esperar, los países de la Unión Europea y en especial los pertenecientes a la OTAN, dieron comienzo también a una serie de reuniones con el objetivo de acordar como actuar e intervenir en el conflicto. Las principales conclusiones de esto fueron dar apoyo a Ucrania y el inicio de una estrategia de boicot económico a Rusia.

Rusia por su parte inició un plan de contramedidas contra las naciones partícipes de este boicot, siendo su principal baza sus recursos energéticos que proveía a muchos de los principales países de Europa.

Estas acciones también serán estudiadas en el siguiente apartado por sus consecuencias económicas, pero en lo referente a lo político, todo este conjunto de sucesos ha ocasionado una enorme inestabilidad en la población y los gobiernos, se ha desarrollado con alarmante rapidez la polarización y radicalización de ideologías

Por otra parte y analizando el ámbito legal, el uso de las redes sociales está principalmente legislado por la Ley Orgánica 2/2018 del 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, así como el Reglamento de la UE 2016/679 del Parlamento Europeo y del Consejo del 27 de abril de 2016.

-Económica: El encarecimiento récord de la energía ha ocasionado conflicto con el uso de aparatos basados en la energía, como los electrodomésticos, **ordenadores y dispositivos móviles.**

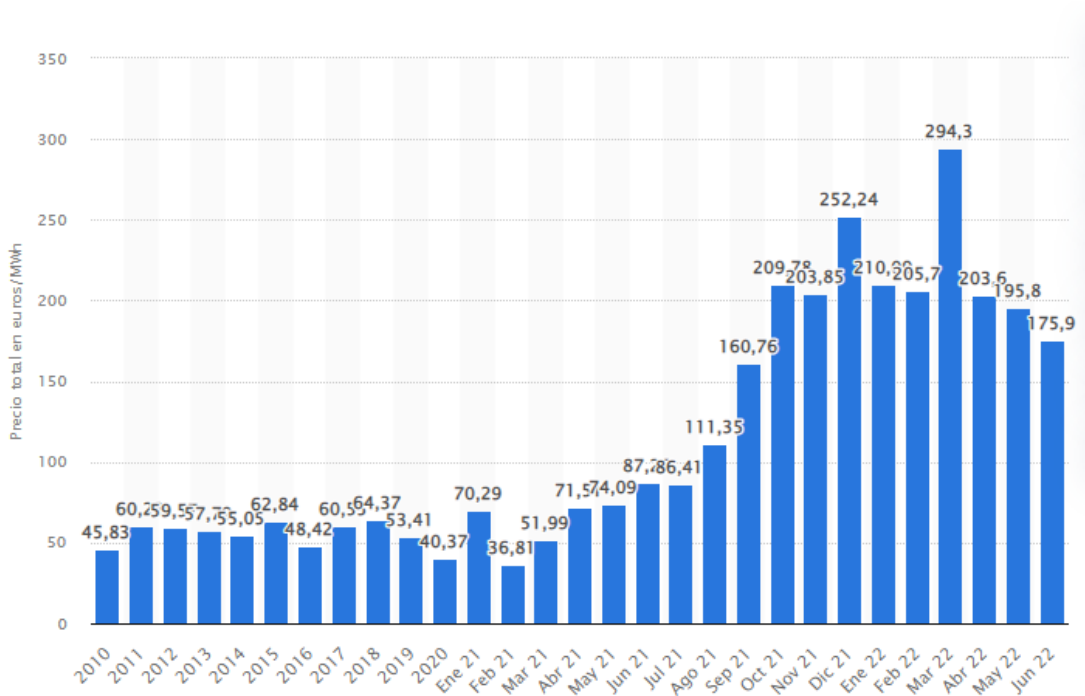


Gráfico 1: Evolución del precio medio final anual de la electricidad en España.

Fuente: Statista, 2022

Por otra parte, todavía no ha finalizado la crisis sanitaria ocasionada por el COVID-19, que trajo un gran parón en la actividad laboral de ininidad de empresas, repercutiendo en los empleos de los ciudadanos españoles desde los inicios de 2020:

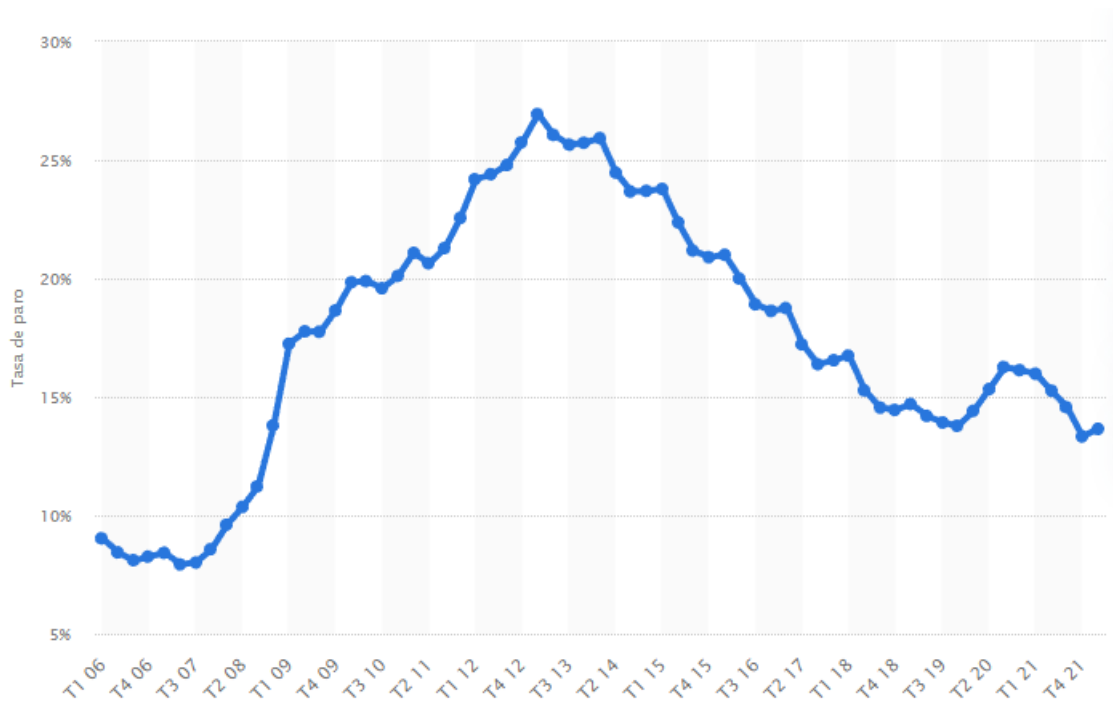


Gráfico 2: Evolución del desempleo por trimestres en España.

Fuente: Statista, 2022

Por tanto, conviene estar muy atento a la evolución del país, pues aunque estaba llegando la recuperación de la crisis sanitaria, si no concluye pronto las hostilidades en Europa podrían acabar siendo una nueva fuente de recesión y recaída en las economías.

-Sociocultural: Desde que se originaron, las redes sociales han sido cada vez más integradas en las vidas de los españoles. Existen actualmente 40,7 millones de usuarios de redes sociales en España, lo que supone un 87,1% de la población.

Además, de media cada usuario utiliza 6 plataformas diferentes de redes sociales, siendo las más populares YouTube, Instagram, Facebook, TikTok y LinkedIn; en ese orden y siendo la primera huésped de 37,4 millones de usuarios.

Por otra parte, entre los hábitos de consumo destacar que de media los españoles gastan unas 6h diarias de media en Internet de las cuales 1 hora y 53 minutos son dedicadas al

uso de redes sociales. De este tiempo invertido, el 71,4% es para información, el 58,6% para entretenimiento y el 38,6% para publicar contenido para otros usuarios.

-Tecnológica: En los inicios de las redes sociales, las principales existentes se utilizaban en ordenadores, pues era el equipo informático más avanzado y en el que solía verse todo el entretenimiento procedente de Internet.

Sin embargo, con la llegada de los smartphone y sus primeros antecesores que permitían el uso de aplicaciones de software que salían de la funcionalidad convencional de un teléfono, el uso mayoritario de las redes sociales migró casi por completo a estos dispositivos, de tal manera que uno de los principales factores determinantes es la accesibilidad a ellas mediante tener o no un smartphone en propiedad.

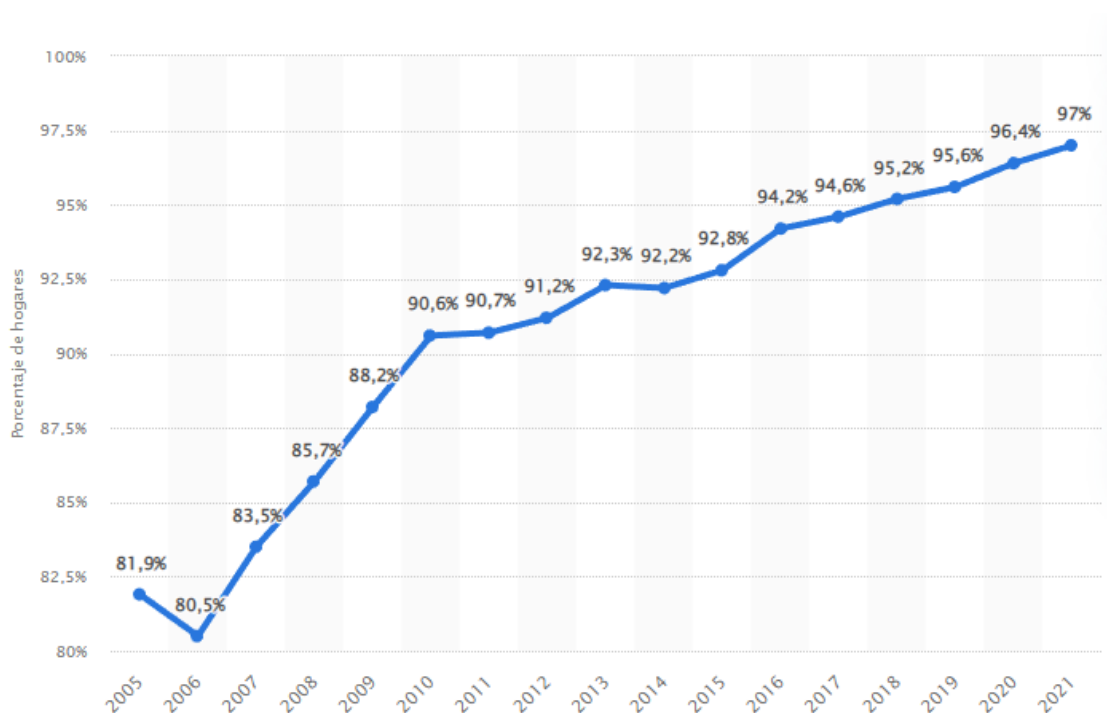


Gráfico 3: Porcentaje de hogares equipados con teléfono móvil en España.

Fuente: Statista, 2021

Como podemos observar, el incremento ha sido estelar, hasta el punto de prácticamente todos los hogares españoles disponen al menos de un dispositivo de estas características.

Además, con los nuevos avances y productos, cada vez más elementos interactúan con las mismas. Son cada vez más las posibilidades por las que las redes sociales se integran

en nuestro día a día, y por tanto la manera en que nosotros nos expresamos a través de ellas.

Con esto finaliza el análisis del macroentorno que afectan al proyecto, por lo que procede recoger las amenazas y oportunidades encontradas en un cuadro resumen para entender a qué problemas se expone y ponen en peligro el proyecto y qué aspectos son aprovechables para impulsar la consecución de los objetivos.

AMENAZAS	OPORTUNIDADES
<u>Inestabilidad política</u>	<u>Inestabilidad política</u>
<u>Crisis energética</u>	<u>Conflicto bélico</u>
<u>Conflicto bélico</u>	<u>Mucho uso de las redes</u>
<u>COVID-19</u>	<u>Accesibles al público</u>
<u>Posible recaída recesión</u>	<u>Buena legislación</u>
<u>Repunte desempleo</u>	<u>RRSS integradas en la vida</u>
<u> </u>	<u>Continuos avances</u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>
<u> </u>	<u> </u>

Cuadro1: Amenazas y oportunidades.

Fuente: Elaboración propia.

Con todo esto, hemos analizado el entorno general y la situación del país dónde va a desarrollarse el proyecto, conociendo así los principales factores que van a ser determinantes para el mismo. Por ello, el siguiente paso es analizar el entorno específico donde va a desarrollarse la actividad.

3.1.2 Microentorno

Para desarrollar un análisis de este tipo, se procede con el uso del modelo de las 5 fuerzas de Porter, que servirá para entender en qué grado podrá tener éxito el proyecto y ser la base del desarrollo de las estrategias de negocio para el mismo.

A. Las 5 fuerzas de Porter

El modelo de las 5 fuerzas de Porter indica la posibilidad de éxito en un sector determinado a través de cinco fuerzas de presión competitiva que pueden clasificarse en:

Fuerzas Horizontales:

- **Competidores actuales:** se trata de un sector donde no existen demasiados competidores en cuanto a número pero sí en cuanto a tamaño y recursos.

Como es de esperar, al ser un ámbito novedoso y que es tendencia en la tecnología, al igual que el resto de tecnologías de machine learning, se encuentran en constante avance y crecimiento tecnológico.

Por otra parte, no existen barreras de entrada y salida severas, pues en general, en el mundo del software cualquiera puede entrar o salir, y las limitaciones presentes se inclinan por dictaminar cuanto tiempo requiere entrar en el sector en función de los recursos disponibles.

- **Competidores potenciales:** El mayor problema de este tipo de barreras del sector reside en que prácticamente cualquier persona o grupo de personas con la voluntad de entrar.

Fuerzas Verticales:

- **Proveedores:** La relación con proveedores en un sector así es un tanto especial. Es cierto que existen cientos de proveedores de software que ofrecen sus módulos y herramientas, pero la relación dentro del desarrollo de aplicaciones suele tender

más a la situación de que el proveedor “lanza” al aire sus herramientas y son los desarrolladores los que deciden cuales eligen y para qué finalidad las van usar.

- **Clientes:** actualmente existen infinidad de clientes potenciales que se encuentran desarrollando tecnología novedosa y experimentando con áreas que podrían beneficiarse del análisis de sentimiento.

Para seguir con el desarrollo de una estrategia, es momento de pasar con el análisis DAFO al completo.

3.2 Análisis interno

En primer lugar, cabe definir las características del proyecto. Se trata de un proyecto actualmente personal, que no cuenta con más recursos adicionales que su creador, con su domicilio y sus equipos informáticos.

Sus puntos fuertes son el entusiasmo, la dedicación y el interés por el sector. Al tratarse de un equipo unipersonal, se garantiza un trato directo y humano.

Por tanto se trata de un proyecto con potencial para, pero que carece de experiencia y recursos propios que lo apoyen.

3.2.1 Análisis funcional

Área comercial	Área de producción
<ul style="list-style-type: none">• Trato personal, cercano y humano• Compromiso y fidelidad• Incapacidad de asumir varios proyectos	<ul style="list-style-type: none">• Escasez de personal• Formación adecuada
Área tecnológica	Área financiera
<ul style="list-style-type: none">• Recursos adecuados para un iniciarse.• Escasez de recursos para proyectos grandes	<ul style="list-style-type: none">• No existen recursos financieros• Necesidad urgente de financiación

Cuadro 2: Resumen análisis funcional.

Fuente: Elaboración propia.

Con este análisis finalizado es posible ahora la creación de un análisis DAFO completo que servirá como guía para la creación de una estrategia.

3.3 Análisis DAFO

Un análisis DAFO consiste en un resumen de los análisis desarrollados en los apartados anteriores, mostrando de manera clara y concisa en un solo golpe de vista las fortalezas y debilidades analizadas internamente así como las amenazas y oportunidades analizadas del entorno exterior que rodea el proyecto.



Cuadro 3: Análisis DAFO.

Fuente: Elaboración propia.

Con esto finaliza por completo el análisis estratégico del proyecto y puede dar comienzo el planteamiento de una estrategia a desarrollar para conseguir los objetivos planteados. Pero antes de dar paso a eso, se incluye un apartado con una simplificada explicación de la parte técnica del proyecto con la que dar a conocer un poco del software en sí y cómo funciona.

3.4 Propósito estratégico

Tras el análisis estratégico y una explicación simplificada del apartado técnico, es el momento de determinar qué estrategia vamos a seguir.

3.4.1 Meta económica y social de la empresa

Este proyecto tiene como metas y objetivos desarrollar un software de análisis de sentimientos en redes sociales y darse a conocer por empresas del sector y el mundo digital.

Estas metas pueden subdividirse en 2 tipos: económicas y sociales.

La meta económica consiste en conseguir financiación y beneficios del software realizado, y tratar de darse a conocer y crecer en cuanto a recursos.

Por otra parte, la meta social definida mejor más adelante con la misión y la visión, es la creación de un equipo de trabajo que se encuentre satisfecho y logre también satisfacer a los clientes.

3.4.2 Misión, visión y objetivos estratégicos

Para definir la estrategia del proyecto, primero hay que definir la misión, visión y objetivos estratégicos.

A. Misión

“Desarrollar software de análisis de sentimiento de calidad y conseguir que el cliente vea satisfechas sus necesidades y resueltos sus objetivos”.

B. Visión

“En 5 años lograr ser conocidos como una empresa de desarrollo de software enfocado al análisis de sentimiento que está iniciándose y ofrece un trato personal y único.”

Se trata de una visión algo optimista pero que es factible llevar a cabo y existen posibilidades de alcanzarla en mayor o menor medida. Para ello es importante definir unos buenos objetivos que guíen el desarrollo y cumplimiento de la misma.

B. Objetivos estratégicos

Los objetivos mencionados para conseguir dicho propósito de manera concisa son:

- **En 1 año:** Lograr un prototipo de software funcional que presente las características base y el potencial de expansión para mostrar ante posibles clientes o inversores.
- **En 2 años:** Puesta en marcha de campaña de publicidad y contacto con clientes potenciales, entrevistas para dar a conocer los servicios ofrecidos y explicar qué necesidades podemos cubrir y aportar soluciones.
- **En 3 años:** Búsqueda de fuentes de financiación e inversores. Contacto con la Generalitat Valenciana para solicitar subvenciones creadas para el desarrollo de PYMES e impulso de empresas tecnológicas. Comienzo de desarrollo de primeros proyectos, evolución de las técnicas empleadas y ampliación de tecnologías conocidas e implementadas
- **En 4 años:** Formación y ampliación de un equipo de trabajo con el que ampliar las capacidades de desarrollo. Puesto que el capital todavía es limitado, búsqueda de perfiles similares, recién egresados que busquen emprender y se les ofrezca la

oportunidad de ser socios mayoritarios del proyecto y de la sociedad en formación.

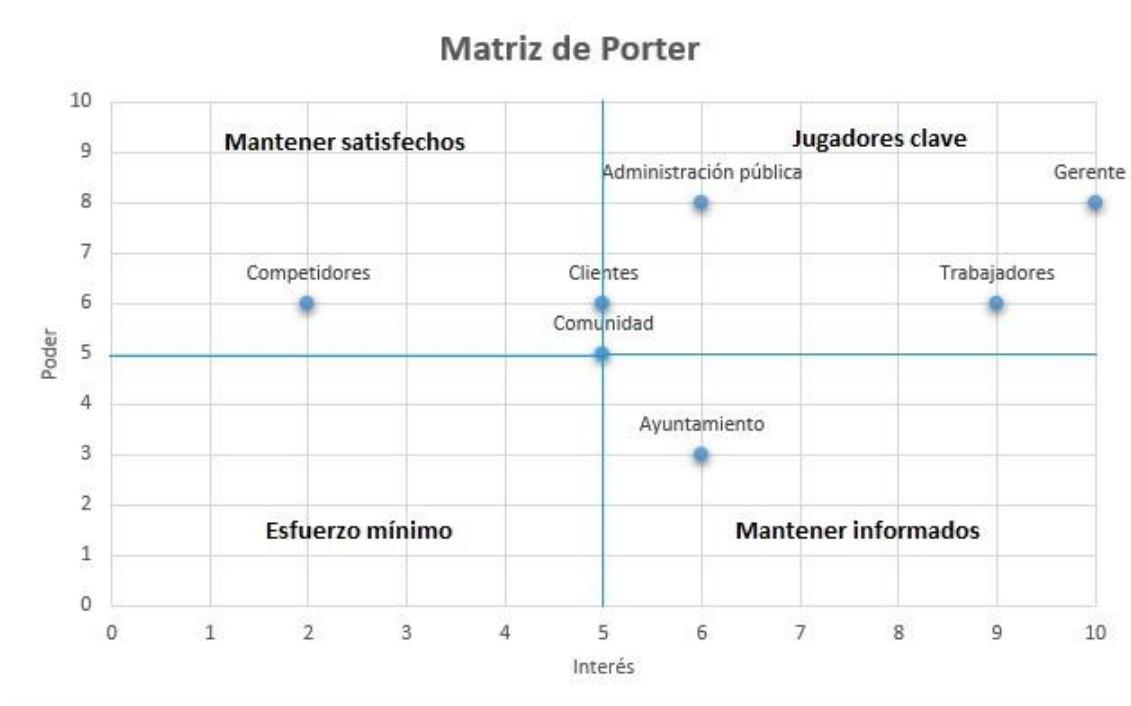
- **En 5 años:** Obtención de primeros resultados de proyectos realizados. Utilizarlos como impulso para llegar a nuevos clientes y ganar reconocimiento, consolidándose como empresa emergente y formal.

Con este propósito estratégico definido por años es el momento de analizar también los grupos de interés relativos al proyecto con el fin de determinar qué sectores deben tener mayor porcentaje de nuestra atención y acciones.

3.4.3 Grupos de interés

Los grupos de interés son aquellos individuos y grupos que están involucrados o interesados en el proyecto, y a causa de ellos van a surgir limitaciones a la hora de realizar los propósitos definidos.

Una vez los tengamos identificados, conviene desarrollar un Mapa de Grupos de Interés con el que posicionar cada grupo en función de su impacto e importancia en la estrategia del proyecto. Con dicho mapa además, se puede apreciar de manera sencilla qué grupos debemos de prestar más atención e interés.



Esquema 1: Matriz de Porter.

Fuente: Elaboración propia.

Con la matriz ya compuesta, es sencillo apreciar cuáles son los grupos a los que debemos prestar especial atención y sobre los que debemos dirigir nuestras acciones.

Con todo esto, tenemos definidas la misión, visión, los objetivos estratégicos y cómo debemos posicionarnos frente a los principales grupos de interés del sector, por lo que toca definir la estrategia a seguir con el proyecto para tratar de garantizar su éxito.

3.5 Estrategia de negocio

Llegó el momento de definir la estrategia a seguir para alcanzar los objetivos planteados. Con ella se busca conducir la compañía hacia la diferenciación de los demás competidores en la búsqueda del éxito. En este proceso van a llevarse diversos pasos.

3.5.1 Creación y mantenimiento de ventajas competitivas

La manera que tienen las empresas de competir entre ellas reside en aquellas características que identifican y caracterizan a una empresa y las diferencian unas de otras, colocándolas así en una posición relativa entre ellas con las que poder competir según la calidad de dichas características.

Puesto que se trata de un proyecto en pleno nacimiento, es prácticamente inviable optar por ventajas competitivas en costes, pues todas las empresas que ya existen con anterioridad en el sector han tenido tiempo de optimizar sus procesos y recursos, por lo que tienen mucho camino recorrido.

Por tanto, tal y como se ha indicado como punto fuerte del proyecto en estos momentos, conviene escoger un enfoque de diferenciación.

Este planteamiento inicial será puesto a prueba en el análisis que se desarrolla a continuación para garantizar que sea la mejor opción para la obtención de una ventaja competitiva o debe replantearse.

3.5.2 La estrategia competitiva

Puesto que este proyecto está todavía comenzando y va a tener dificultades para competir en el sector, tratar de encontrar un nicho que atender puede ser la solución más factible al principio, ofreciendo una diferenciación notoria respecto al resto de competidores.

Un nicho ideal que ya ha sido analizado con anterioridad es la Administración pública, pues al no ser una empresa privada sus objetivos no residen en generar beneficios, sino que están orientados a obtener beneficios sociales para sus ciudadanos. Por ello, y dada la naturaleza del proyecto, puede ajustarse perfectamente y ofrecer un servicio que dicha administración podría utilizar para fomentar el cuidado de la salud mental y seguir con los objetivos de desarrollo sostenible, así como promover el uso responsable de las redes sociales. Además, puesto que es la misma entidad que asigna subvenciones para el desarrollo de PYMES y desarrollo tecnológico, es una buena manera de obtener la financiación que se requiere para crecer.

3.5.3 Estrategia en función del ciclo de vida

Para reforzar la elección de una estrategia competitiva hay que tener en cuenta el ciclo de vida del proyecto.

Cada etapa viene definida por una serie de características que indicaran qué acciones son recomendables tomar en cada una para tratar de garantizar el éxito

En nuestro caso, el proyecto se encuentra en la etapa de **Introducción**, por lo que es de esperar que no reporte muchos beneficios y es prioridad darle recursos a una buena campaña que le aporte visibilidad.

3.6 Estrategia corporativa

En este apartado se va a desarrollar la estrategia corporativa que seguirá la empresa organización en formación. Con ella se van a tomar decisiones a nivel estratégico que implicarán en la estructura organizativa futura y el resto de nuevas estrategias posteriores.

3.6.1 Direcciones de desarrollo

Como metodología para trabajar este concepto, se prepara la conocida Matriz de Ansoff:

De entre todas las opciones, puesto que este proyecto está comenzando, la organización es nueva y sí o sí se va a tratar de un nuevo mercado para la organización y un nuevo producto, por lo que estamos hablando de una estrategia de diversificación.

La estrategia de diversificación es la más drástica y arriesgada, pues implica innovar y entrar en entornos competitivos nuevos para la empresa. A la hora de diversificar, se puede optar por dos variantes:

- **Diversificación relacionada**
- **Diversificación no relacionada**

Una vez más, ya que no existen precedentes, la opción escogida es la **diversificación no relacionada**.

En nuestro caso invertiremos en un nuevo sector emergente y asumiremos la mayoría de riesgos de esta opción, siendo el mayor de ellos la falta de experiencia y habilidades dentro de dicho sector, que se irá paliando con el tiempo.

3.6.2 Métodos de desarrollo

Para desarrollar la dirección escogida, puede optarse por dos opciones de manera general: el desarrollo interno y el desarrollo externo.

Para nuestro proyecto, de manera obvia, la opción a elegir es el desarrollo interno, pues al estar todavía arrancando y dando comienzo, hay que construir los cimientos en primer lugar. Obtención de financiación, añadir recursos a los existentes, contratar más personal serán acciones que definiremos en la implantación de la estrategia.

4.PROPUESTA DEL SOFTWARE

4.1 Visión general

El objetivo de este apartado es aportar una visión general del proyecto software a desarrollar con la que asentar los conceptos más importantes de la implementación.

El primer paso es la obtención de una fuente de texto con la que trabajar. Puede extraerse de bases de datos de mensajes recogidos, pero en nuestro caso lo haremos en vivo de la propia red social.

Una vez tengamos extraída la materia bruta del software, hay que procesarla para que sea cómodo trabajar con ella. Esto ocurre porque como es tendencia en las redes sociales actuales, a la hora de realizar publicaciones en el propio texto se insertan herramientas de la red para interactuar con sus sistemas, funciones y resto de usuarios. Entre estos elementos se encuentran hashtags, menciones, retweets, otras publicaciones, fotos, vídeos, enlaces y otras muchas más de índole similar. Para limpiar el texto, tendremos que desarrollar código capaz de analizar la secuencia completa y eliminar dichos elementos, dejando sólo el mensaje original que desea transmitir su emisor, el que podría comunicarse en una conversación en persona.

Con los mensajes listo, procede su análisis y categorización, someterlos a las funciones de análisis de sentimiento oportunas para que ejerzan un veredicto acerca de la intencionalidad del texto en función de qué expresiones se han elegido para transmitirlo. Tras juzgar su intencionalidad, se le asignará a cada elemento un valor o característica que lo defina como Positivo o Negativo, categorizándolo dentro de una de las etiquetas predefinidas por el modelo.

Por último, como es de esperar deberemos mostrar los resultados obtenidos, por lo que mostraremos un análisis individual de cada mensaje con su valoración y categorización y los resultados de un análisis en conjunto de todos los mensajes estudiados en la muestra.

4.2 Metodología

Para el desarrollo de software existen diversas metodologías para elaborar los proyectos software. A lo largo de la carrera hemos estudiado varias, siendo las dos más indagadas:

- **Metodología Scrum:** Se trata de una metodología ágil basada en los denominados Sprints. Cada Sprint consiste en un periodo corto de tiempo, normalmente un par de semanas, en las que el equipo trabaja por tareas repartidas en ir implementando aspectos del proyecto software. La idea es que sea un desarrollo incremental de mejora, y se prioriza la funcionalidad desde día 1 y que esta con cada Sprint vaya ganando nuevas funcionalidades y añadidos. Además se definen una serie de roles con responsabilidades para el desarrollo del proceso.

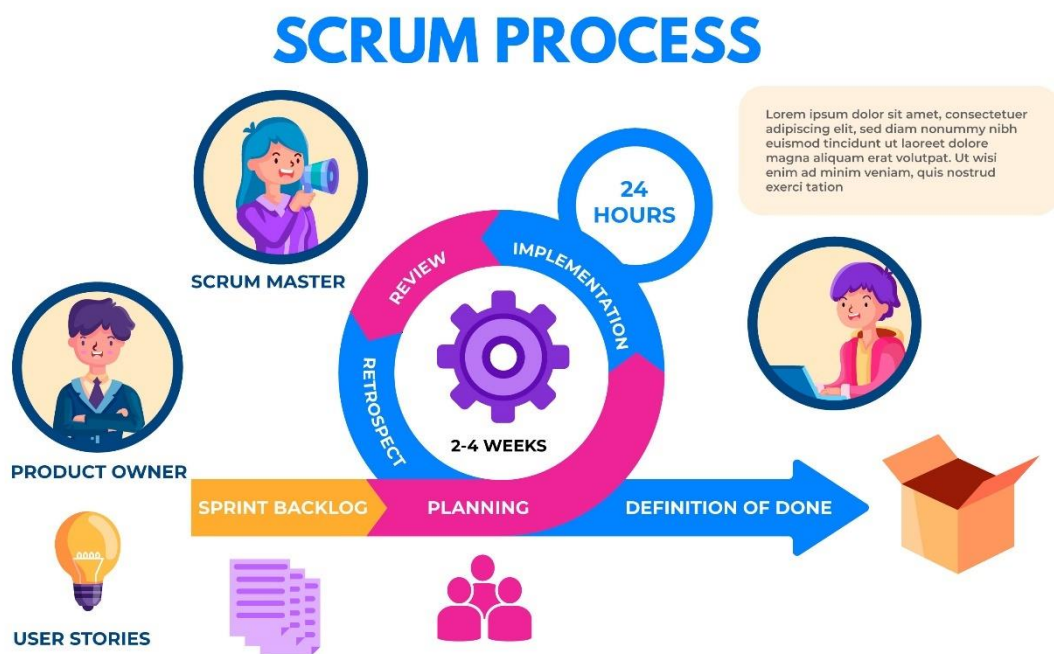


Imagen 1 :Metodología Scrum.

Fuente: Google imágenes.

- **Metodología en Cascada:** Es una de las más tradicionales de desarrollo. El proyecto se subdivide en etapas que forman un avance secuencial de cara a una revisión final al terminar el proceso. Normalmente no existe una muestra funcional hasta las fases más finales del desarrollo.

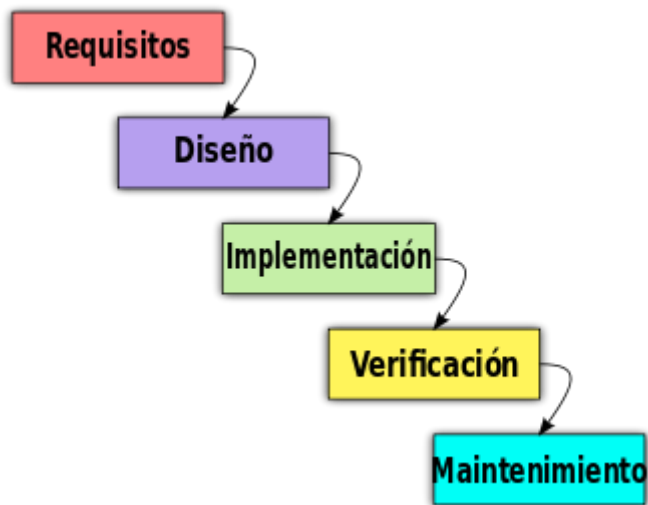


Imagen 2 :Metodología en cascada.

Fuente: Google imágenes.

Para el desarrollo de este software, puesto que gran parte de su complejidad reside en el conocimiento detrás de su funcionamiento y no en el desarrollo de código, vamos a optar por la metodología en cascada, pues no requiere de tantos sprints y retroalimentación constante para su elaboración

4.3 Requisitos del sistema

Siguiendo con la metodología planteada y con la visión general desarrollada, hay que especificar un poco más y desarrollar un listado de requisitos a cumplir por el software teniendo en cuenta el listado de necesidades y limitaciones a ofrecer.

4.3.1 Alcance del proyecto

Con la visión general desarrollada, hay que especificar un poco más y desarrollar un listado de requisitos a cumplir por el software teniendo en cuenta el listado de necesidades y limitaciones a ofrecer.

Este trabajo tiene como finalidad analizar, diseñar, implementar y probar un sistema capaz de extraer mensajes ofrecidos por usuarios de Twitter y analizar la información subjetiva que aportan con ellos para categorizarlos como positivos o negativos.

4.3.2 Requisitos funcionales

Sistema

RF-1	El sistema será un software desarrollado para extraer información y análisis de mensajes de Twitter
RF-2	El sistema permitirá reconocer la información subjetiva transmitida en cada mensaje para conocer su intencionalidad

Autenticación

RF-3	Para el uso de la aplicación deberán requerirse claves de acceso especiales entregadas a cuentas de desarrollador de Twitter
-------------	--

Extracción de información

RF-4	El sistema permitirá la búsqueda de información asociada a un hashtag determinado
RF-5	El sistema permitirá la búsqueda de información asociada a un nombre de usuario determinado

4.3.3 Requisitos no funcionales

Documentación

RNF-1	El código debe tener documentadas tantas anotaciones de código como se requiera para su correcta comprensión y mantenimiento.
--------------	---

Seguridad

RF-2	El software sólo debe funcionar con las credenciales correctas de una cuenta de desarrollador de Twitter tal y como se especifica en los acuerdos de la red social.
-------------	---

Usabilidad

RNF-3	Los requisitos establecidos en cuanto a datos finales deben ser capaces de ser mostrados
RNF-4	El software debe ser fácil de comprender en cuanto a funcionamiento para su correcta comprensión e integración en otros sistemas

Eficiencia

RNF-5	La aplicación debe funcionar de manera eficiente y ejecutarse dentro de márgenes de tiempo de entre 15 y 20 segundos
RNF-6	La aplicación debe ser ágil y poseer un rendimiento óptimo para su uso

Mantenibilidad

RNF-7	La aplicación debe garantizar su estabilidad integra
RNF-8	La aplicación debe ser fácil de analizar en cuanto a errores de funcionamiento
RNF-9	El software debe presentar facilidad a la hora de realizarse cambios dentro del mismo
RNF-10	La aplicación debe ser fácil de someter a pruebas que la evalúen

Una vez finalizada la propuesta del software procedemos a revisar las tecnologías relacionadas con el mismo.

5.REVISIÓN DE LAS TECNOLOGÍAS Y MEDIOS UTILIZADOS

En este apartado se va a desarrollar y listar las posibilidades de entornos, redes, lenguajes y en definitiva tecnología considerada para el proyecto, así como su elección y su explicación.

5.1 Redes Sociales

Existen multitud de redes sociales con las que expresarse, de entre ellas, vamos a explicar las características de algunas de las más famosas:

- **Facebook:** Una de las más veteranas y pioneras en el sector. Se trata de una red social bastante completa que abarca multitud de usos y funcionalidades. En ella el usuario tiene un panel personal donde puede acceder y visualizar páginas creadas, ver publicaciones de otros usuarios, gestionar invitaciones a eventos, comenzar chats, conocer lugares y personas y un largo etcétera.

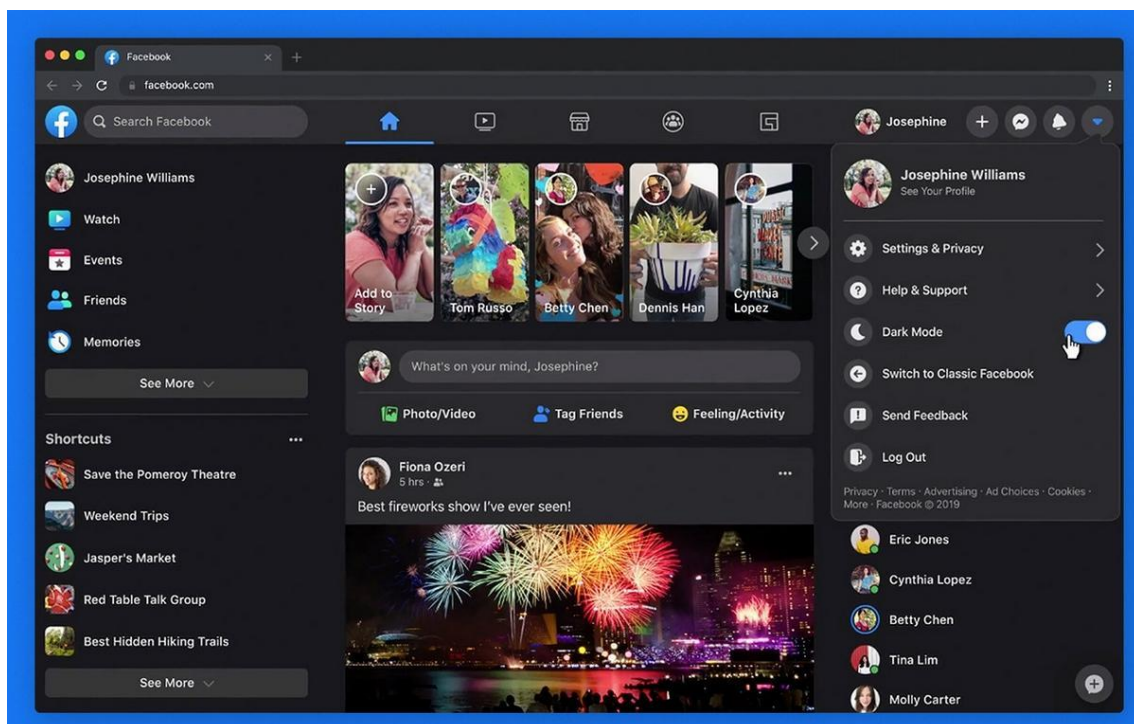


Imagen 3 :Interfaz de ejemplo de Facebook.

Fuente: Google imágenes.

En definitiva y como se puede observar, es una red social muy **completa** que abarca muchos formatos de uso: fotos, vídeos, texto, páginas...

- **Instagram:** Una de las redes sociales más utilizadas por los jóvenes y reciente adquisición de la compañía anterior, Facebook (ahora conocida como Meta). Su principal atractivo reside en la facilidad de uso y de consumo, pues se fundamenta principalmente en fotos y vídeos cortos, que van apareciendo uno tras otro con un simple scroll hacia abajo o de manera continuada si se trata de historias. Además repunta por su popularidad líder en uso desde smartphome.



Imagen 4 :Interfaz maqueta de Instagram.

Fuente: dreamstime.com

- **Twitter:** La red social de las opiniones por excelencia. Se fundamenta en un sistema en el que cada usuario tiene la opción de comentar cualquier cosa que le esté ocurriendo en 140 caracteres. Tiene además soporte para complementar con fotos, vídeos y enlaces, pero la base reside ahí. Cada opinión o comentario es llamado *tweet* y el sistema lo recoge y almacena, para ir mostrándolos en lista uno debajo de otro ordenados cronológicamente a cada usuario según sus preferencias, intereses y personas a las que sigue.



Imagen 5 :Interfaz de ejemplo de Twitter.

Fuente: Google imágenes

Además, tal y como comentaremos más adelante, es una de las redes sociales más amigable con los desarrolladores externos, pues gran parte de su código es abierto al público y ella misma ofrece cuentas de desarrolladores para acceder al uso de APIs oficiales con las que trabajar.

De entre las 3 opciones, por su naturaleza y comodidad para trabajar con su formato, es escogida Twitter desde el inicio del proyecto, pues trabajar con texto es más asequible para el análisis de sentimiento y la propia red almacena los tweets de manera organizada y cómoda para trabajar en bases de datos. Las otras redes sociales son estudiadas para una posible extensión o migración a ellas.

5.2 Tecnologías y entornos

Una vez escogida la red social, cabe elegir el lenguaje de programación que va a utilizarse para el desarrollo del código. Entre las opciones más representativas para el machine learning tenemos Java y Python.

Por una parte, Java es uno de los lenguajes concurrentes más utilizados a nivel mundial para el desarrollo de todo tipo de software. Lleva muchos años en el sector, por lo que es de esperar que existan librerías y recursos para prácticamente cualquier problemática software que se haya planteado hasta la fecha.

Entre sus librerías, se encuentra *Twitter4J*, una librería que permite la integración de la API oficial de Twitter, permitiendo así la autenticación, obtención y parametrización de datos directamente de la propia Twitter con tan solo aportar los credenciales de una cuenta de desarrollador. Además destacar que para ciertos planteamientos más complejos, Java presenta muchas facilidades de integración con las principales bases de datos del mercado, como MySQL.

Por otra parte, tenemos Python, uno de los lenguajes que más popularidad ha ganado durante los últimos años debido a su gran facilidad de uso y aprendizaje amigable con los nuevos usuarios. Es menos estricto y estructurado que Java, por lo que presenta una mayor sencillez de uso a costo de no alcanzar tantas funcionalidades complejas como él.

Entre sus librerías más conocidas para el ámbito, se encuentra *Tweepy* con la que podemos obtener mensajes directamente de Twitter en tiempo real de manera similar a su hermana de Java, también aportando los credenciales de autenticación de una cuenta developer de Twitter.

Entre estas dos opciones, nos decantaremos por Python por varios motivos: su sencillez de uso, atender la recomendación personal de mi profesor de Estructura de Datos y

Algoritmos tras la consulta técnica que le realicé mientras ideaba el proyecto y la más subjetiva de todas, por variar. A lo largo de estos años de carrera he estudiado y puesto en práctica mucho Java en comparación a otros lenguajes, por lo que parte de la motivación reside en el reto de cambiar de entorno de confort.

5.3 Elección de un modelo basado en Transformers pre entrenado

En este proyecto, de entre las diversas opciones de análisis de sentimiento existentes y revisadas anteriormente, se escoge la utilización de un modelo basado en transformers pre entrenado, específicamente el modelo BERT. La elección de este enfoque se basa en varias ventajas significativas que ofrece.

Una de las principales fortalezas de BERT radica en su capacidad para capturar el contexto bidireccional en el texto. A diferencia de los modelos secuenciales tradicionales, BERT considera tanto las palabras anteriores como las posteriores a cada palabra en una oración. Esto es crucial en el análisis de sentimiento, donde el significado y la polaridad de una palabra pueden estar influenciados por el contexto circundante. Al capturar las relaciones contextuales complejas, BERT mejora la comprensión del texto y la precisión en la identificación del sentimiento expresado.

Otra ventaja clave de BERT es su habilidad para generar representaciones contextualizadas de las palabras. Durante el preentrenamiento, BERT aprende a asignar pesos a las palabras en función del contexto en el que aparecen. Esto permite capturar mejor la semántica y las relaciones entre las palabras, lo cual es esencial en el análisis de sentimiento. Al utilizar representaciones contextualizadas, BERT puede interpretar adecuadamente las palabras en función de su contexto, mejorando la precisión en la clasificación de sentimientos en los tweets.

El preentrenamiento de BERT en grandes cantidades de datos sin etiquetar también juega un papel fundamental en su efectividad. Durante este proceso, BERT adquiere conocimientos generales sobre el lenguaje y patrones lingüísticos. Esta capacidad de

generalización es especialmente relevante en el análisis de sentimiento en Twitter, donde los textos pueden ser informales, con jerga y emoticonos. Al aprovechar el preentrenamiento, BERT puede utilizar su conocimiento previo para comprender y clasificar adecuadamente los sentimientos expresados en los tweets.

Por último, BERT se destaca por su adaptabilidad a dominios específicos. A través del ajuste fino, se puede personalizar el modelo para que se adapte mejor a las peculiaridades y características propias de Twitter. Mediante el entrenamiento con datos etiquetados en la tarea de análisis de sentimiento, BERT puede mejorar su capacidad de clasificación y proporcionar resultados más precisos y comprensivos en el contexto de Twitter.

6.PLANTEAMIENTO DE LA SOLUCIÓN

Con las tecnologías y entornos ya decididos, llega el momento de entrar en materia y comenzar a desarrollar una solución a la problemática planteada.

6.1 Planteamiento teórico

En primer lugar, antes de comenzar a escribir código, hay que plantear una guía de pasos a realizar a nivel conceptual para llevar un desarrollo coherente.

Nuestro objetivo con este software es extraer un conjunto de información de Twitter, procesarlo y extraer una serie de resultados mostrables.

Para ello, vamos a seguir el siguiente esquema de hitos:

1. Importación de librerías
2. Conectar con la API de Twitter
3. Extraer el contenido deseado
4. Limpiar y preparar el contenido para su procesamiento
5. Procesamiento y categorización de la información: asignación de valores según su análisis
6. Agrupación de los datos según coincidencia
7. Construcción de elementos informativos
8. Representación de los resultados por pantalla

Para acompañar este esquema, conviene realizar un diagrama de flujo que exprese de manera visual la intencionalidad y funcionalidad del mismo.

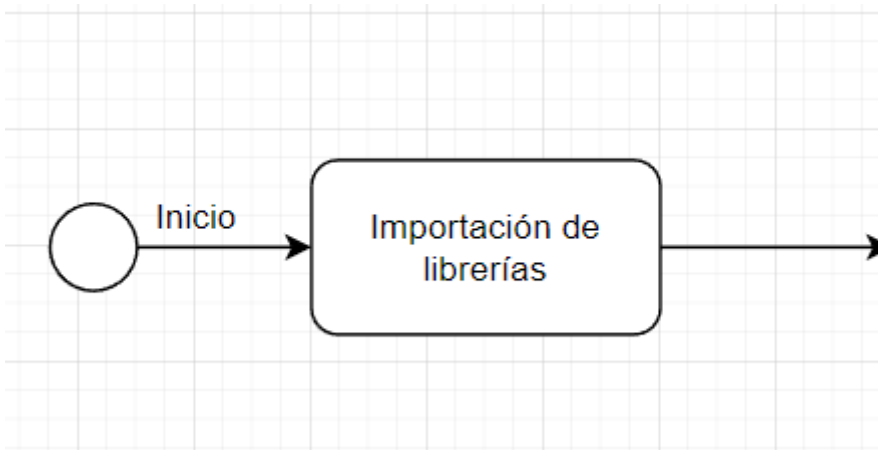


Diagrama 1 :Importación de librerías.

Fuente: Elaboración propia

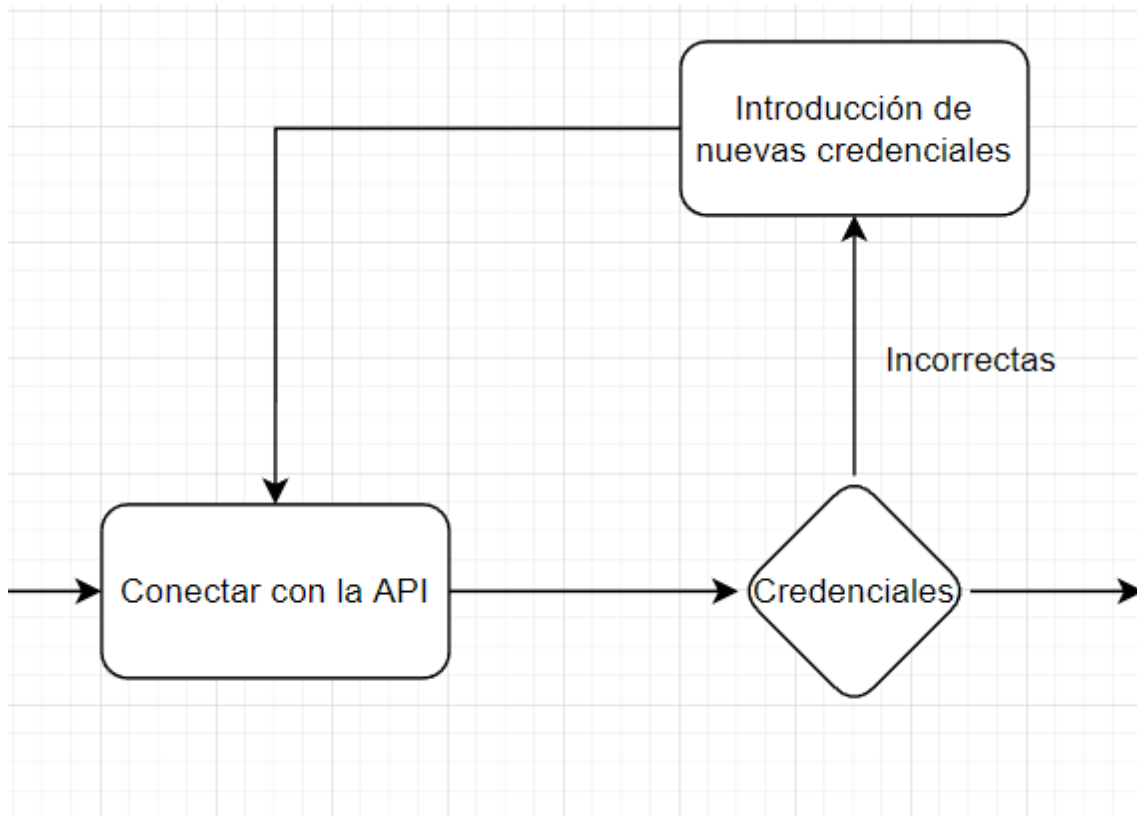


Diagrama 2 :Autenticación.

Fuente: Elaboración propia

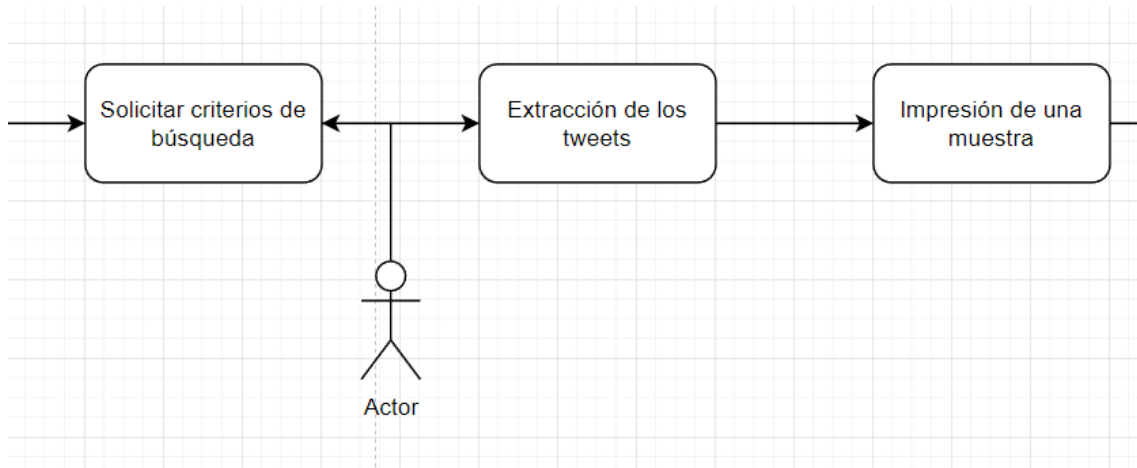


Diagrama 3 Búsqueda.

Fuente: Elaboración propia

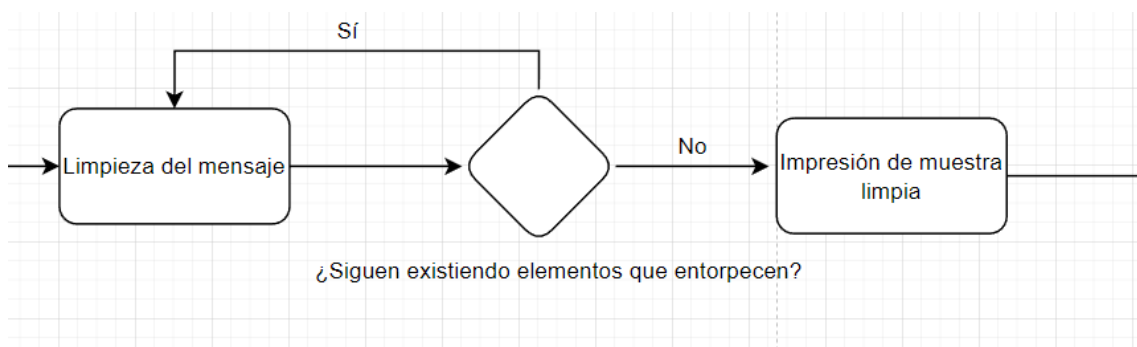


Diagrama 4 :Limpieza del mensaje.

Fuente: Elaboración propia

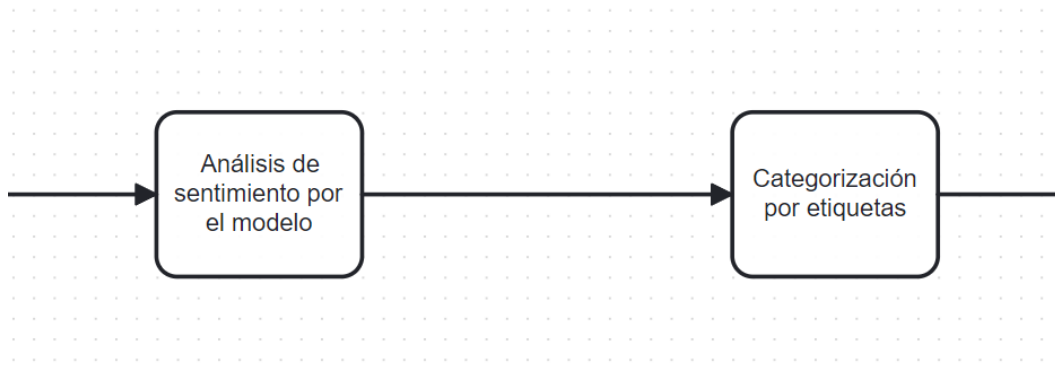


Diagrama 5 :Asignación y agrupación.

Fuente: Elaboración propia

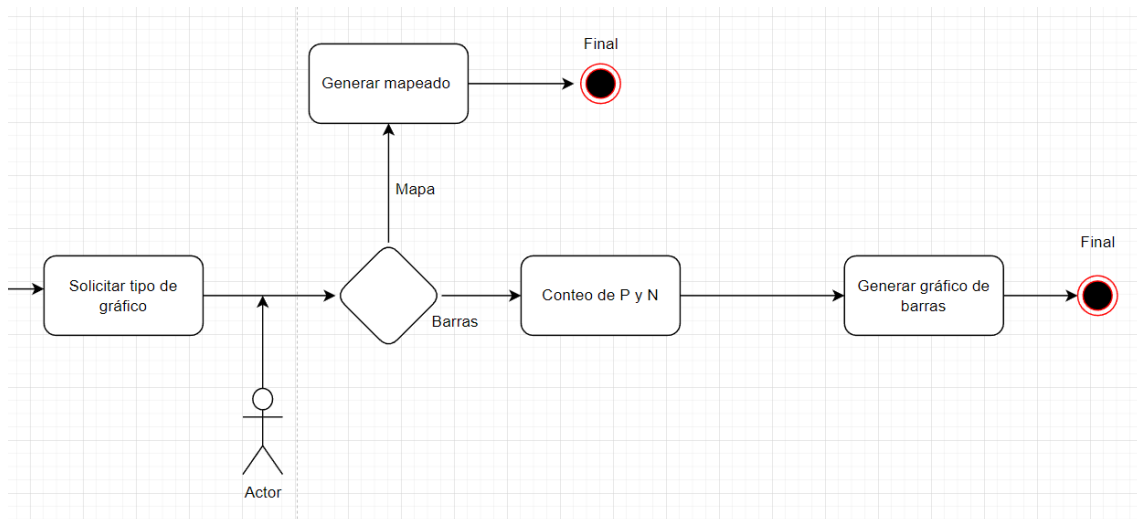


Diagrama 6 :Muestreo de resultados.

Fuente: Elaboración propia

Con el planteamiento claro, podemos proceder con el comienzo del desarrollo de código.

6.2 Desarrollo de código

En este apartado vamos a desarrollar las principales partes del código que tienen una relevancia significativa para el mismo.

6.2.1 Importación de librerías

```
import tweepy
from transformers import pipeline
import re
import matplotlib.pyplot as plt
```

Código 1 : Importación de librerías.

Fuente: Elaboración propia y repositorios de GitLab

En primer lugar, como es de esperar en cada desarrollo software precede la importación de todas las herramientas que utilizará el programa.

La librería tweepy es una herramienta de Python que facilita la interacción con la API de Twitter. Proporciona métodos y clases para autenticarse, acceder a la información del usuario y realizar acciones como buscar tweets, publicar tweets, entre otros.

La librería transformers es una biblioteca de Python que ofrece una interfaz para utilizar modelos de aprendizaje automático de lenguaje natural (NLP). La utilizamos para acceder a un modelo pre entrenado de análisis de sentimientos.

La librería re es el módulo de expresiones regulares de Python. Se utiliza para realizar operaciones de búsqueda y manipulación de cadenas de texto mediante patrones.

La librería matplotlib.pyplot es una biblioteca de visualización en Python que permite crear gráficos y visualizaciones. Se utiliza para representar los resultados del análisis de sentimientos en forma de gráficos.

6.2.2 Conexión con la API de Twitter

Para la conexión con la API, tal y como se ha comentado anteriormente, hemos creado una cuenta de desarrollador en Twitter. Con ella, hemos creado la aplicación en sus servidores para que nos generen una serie de identificadores para la misma.

```
# Configuración de las credenciales de Twitter
consumer_key = [REDACTED]
consumer_secret = [REDACTED]
access_token = [REDACTED]
access_token_secret = [REDACTED]

# Configuración de las credenciales de la API de Twitter
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

# Crear una instancia de la API de Twitter
api = tweepy.API(auth)
```

Código 2 :Conexión con la API.

Fuente: Elaboración propia y repositorios de GitLab

Estos identificadores se corresponden por orden de aparición:

- A la clave API que se asocia a la aplicación, permitiendo la identificación del cliente.
- La contraseña del cliente para establecer la autenticación
- El token de acceso que utiliza los parámetros anteriores para establecer el nivel de privilegio de acceso, es decir, determinar a qué datos se pueden acceder y a cuáles no.
- La clave de seguridad del propio token

Por razones obvias de confidencialidad y seguridad, en la captura no se muestran los identificadores.

Una vez configuradas las credenciales, se crea una instancia de la API de Twitter llamada `api` utilizando el objeto `auth` como parámetro. Esta instancia de la API permitirá interactuar con la API de Twitter y

realizar diversas operaciones, como obtener tweets, publicar tweets, buscar tweets, entre otras.

6.2.3 Extracción de contenido

```
# Pedir al usuario que ingrese el nickname de Twitter
nickname = input("Ingresar el nickname de Twitter: ")

# Pedir al usuario que ingrese el número de tweets a analizar
num_tweets = int(input("Ingresar el número de tweets a analizar: "))
```

Código 3 :Solicitud de parámetros

Fuente: Elaboración propia y repositorios de GitLab

```
# Obtener los últimos tweets del usuario
tweets = api.user_timeline(screen_name=nickname, count=num_tweets, tweet_mode="extended")
```

Código 4 :Extracción por nick.

Fuente: Elaboración propia y repositorios de GitLab

Después de obtener el nickname de Twitter y el número de tweets a analizar, se utiliza la instancia de la API de Twitter (api) para llamar al método user_timeline. Este método recibe como argumentos el screen_name (nickname de Twitter), el count (número de tweets a obtener) y el tweet_mode establecido en "extended" para obtener el texto completo de los tweets.

La función user_timeline devuelve una lista de objetos de tweet que cumplen con los criterios especificados. Esta lista de tweets se guarda en la variable tweets y se utilizará más adelante para realizar el análisis de sentimientos de cada tweet.

6.2.4 Clasificador basado en transformers

```
# Crear el clasificador de sentimientos utilizando Transformers con un modelo preentrenado
classifier = pipeline("sentiment-analysis", model="nlpTown/bert-base-multilingual-uncased-sentiment")
```

Código 5 :Clasificador

Fuente: Elaboración propia y repositorios de GitLab

En esta parte del código se crea un clasificador de sentimientos utilizando la biblioteca Transformers. Se utiliza un modelo preentrenado llamado "nlpTown/bert-base-multilingual-uncased-sentiment".

La función pipeline de Transformers permite construir un pipeline de procesamiento de texto que incluye diferentes tareas, como el análisis de sentimientos. En este caso, se utiliza el pipeline de "sentiment-analysis" para realizar el análisis de sentimientos.

El argumento "model" especifica el nombre o la ruta del modelo preentrenado que se desea utilizar. En este caso, se utiliza el modelo "nlpTown/bert-base-multilingual-uncased-sentiment", que es una variante del modelo BERT entrenado en varios idiomas.

Al utilizar el modelo preentrenado en el clasificador de sentimientos, se aprovechan los conocimientos y la capacidad de generalización que el modelo ha aprendido durante su entrenamiento en grandes conjuntos de datos. Esto permite que el clasificador pueda determinar el sentimiento de un texto dado, asignando una etiqueta de sentimiento y un puntaje de confianza que indica la probabilidad asociada a esa etiqueta.

Estas etiquetas son:

[1 Star] Muy Negativo

[2 Star] Negativo

[3 Star] Neutro

[4 Star] Positivo

[5 Star] Muy Positivo

Por último, al crear el clasificador de sentimientos, se asigna a la variable classifier. Este objeto classifier se utilizará más adelante para el análisis.

6.2.5 Análisis de sentimientos

```
# Realizar la evaluación individual de cada tweet
tweet_labels = []
tweet_scores = []
for tweet in tweets:
    if "retweeted_status" in dir(tweet): # Ignorar retweets
        continue
    # Limpiar el texto del tweet
    clean_text = re.sub(r"http[S+]\.pic.twitter.com\[S+]\@[S+]\#[S+]", "", tweet.full_text)

    # Realizar el análisis de sentimientos en el texto del tweet
    result = classifier(clean_text)[0]

    # Guardar el label y el score del tweet
    tweet_labels.append(result["label"])
    tweet_scores.append(result["score"])

    # Mostrar el tweet y su evaluación
    print("Tweet:", tweet.full_text)
    print("Label:", result["label"])
    print("Score:", result["score"])
    print("==" * 50)

# Concatenar los textos de los tweets en un solo string
text = " ".join([re.sub(r"http[S+]\.pic.twitter.com\[S+]\@[S+]\#[S+]", "", tweet.full_text) for tweet in tweets if
                 "retweeted_status" not in dir(tweet)])

# Realizar la evaluación global del conjunto de tweets
results = classifier(text)
global_labels = [result["label"] for result in results]
global_scores = [result["score"] for result in results]
```

Código 6 :Análisis de sentimientos

Fuente: Elaboración propia y repositorios de GitLab

Primero, se crean dos listas vacías llamadas tweet_labels y tweet_scores. Estas listas se utilizarán para almacenar las etiquetas de sentimiento y los puntajes correspondientes de cada tweet analizado.

A continuación, se itera sobre cada tweet en la lista tweets obtenida previamente. Se verifica si el tweet es un retweet utilizando la condición `if "retweeted_status" in dir(tweet)`. En caso de ser un retweet, se omite y se pasa al siguiente tweet.

Luego, se procede a limpiar el texto del tweet mediante el uso de expresiones regulares. Se eliminan las URLs, las menciones de usuario y los hashtags. El texto limpio se almacena en la variable `clean_text`.

A continuación, se realiza el análisis de sentimientos utilizando el clasificador previamente creado, que se inicializó con el modelo pre entrenado. Se pasa el texto limpio del tweet como entrada al clasificador y se obtiene el resultado en forma de lista.

El resultado obtenido es un diccionario que contiene la etiqueta de sentimiento y el puntaje asociado al sentimiento más probable. Se toma el primer elemento de la lista (`[0]`) para obtener la etiqueta y el puntaje del sentimiento más probable.

La etiqueta y el puntaje del sentimiento se agregan a las listas `tweet_labels` y `tweet_scores`, respectivamente, para su posterior uso.

Además, se muestra por pantalla el texto del tweet, la etiqueta de sentimiento y el puntaje asociado al sentimiento para cada tweet analizado. Esto permite tener una visión individual de cada tweet y su evaluación de sentimiento.

Posteriormente, se concatena el texto de todos los tweets en un solo string utilizando comprensión de listas y se guarda en la variable `text`. Esto es necesario para realizar la evaluación global del conjunto de tweets.

Para la evaluación global, se llama al clasificador con el texto concatenado (text) como entrada. El resultado es una lista de diccionarios que contienen las etiquetas de sentimiento y los puntajes asociados al sentimiento más probable para el texto completo.

Finalmente, se extraen las etiquetas y los puntajes de sentimiento de los resultados globales y se almacenan en las listas global_labels y global_scores, respectivamente. Estas listas representan la evaluación global del conjunto de tweets.

6.2.6 Construcción de elementos informativos

Ya casi finalizando, con la información ya trabajada y agrupada, es el momento de generar un contenido visual que muestre los resultados de manera intuitiva:

```
# Crear una figura y ejes para la visualización gráfica
fig, (ax1, ax2, ax3) = plt.subplots(3, 1, figsize=(8, 9))

# Graficar los resultados individuales de cada tweet
ax1.bar(range(len(tweet_labels)), tweet_scores)
ax1.set_xticks(range(len(tweet_labels)))
ax1.set_xtickLabels([f"Twit {i+1}\n{label}" for i, label in enumerate(tweet_labels)], rotation=45)
ax1.set_ylabel("Score")
ax1.set_title("Sentimientos de cada tweet")

# Graficar los resultados globales del conjunto de tweets
ax2.bar(range(len(global_labels)), global_scores)
ax2.set_xticks(range(len(global_labels)))
ax2.set_xtickLabels(global_labels, rotation=45)
ax2.set_ylabel("Score")
ax2.set_title("Sentimientos del conjunto de tweets")

# Contar los tweets por cada label
label_count = {}
for label in tweet_labels:
    label_count[label] = label_count.get(label, 0) + 1

# Obtener todas las labels y sus conteos, incluso si tienen 0 elementos
all_labels = list(set(tweet_labels))
all_labels.sort()
counts = [label_count.get(label, 0) for label in all_labels]

# Graficar la cantidad de tweets por cada label
ax3.bar(all_labels, counts)
ax3.set_ylabel("Cantidad")
ax3.set_title("Cantidad de tweets por label")

# Ajustar la distribución de los subplots y mostrar la gráfica
plt.tight_layout()
plt.show()

# Guardar como fichero
plt.savefig("C:/Users/soler/Desktop/resultados.png")
```

Código 7: Generación de gráficos

Fuente: Elaboración propia y repositorios de GitLab

En primer lugar, se utiliza el objeto `ax2` para crear un gráfico de barras. Se establece el rango de los valores en el eje x utilizando `range(len(global_labels))`. Luego, se establecen los ticks en el eje x y se asignan las etiquetas correspondientes a partir de `global_labels`. El parámetro `rotation=45` indica que se rotarán las etiquetas para una mejor legibilidad. Se establecen los ejes y títulos de las etiquetas y puntajes de sentimiento, y se asigna el título "Sentimientos del conjunto de tweets" al gráfico.

En segundo lugar, se crea un diccionario llamado `label_count` para contar la cantidad de tweets por cada etiqueta de sentimiento. Se itera sobre la lista `tweet_labels` y se actualiza el conteo para cada etiqueta en el diccionario utilizando `label_count.get(label, 0) + 1`. Si la etiqueta no existe en el diccionario, se inicializa con 0 y luego se incrementa en 1.

A continuación, se crea una lista llamada `all_labels` que contiene todas las etiquetas únicas presentes en `tweet_labels`. Las etiquetas se ordenan alfabéticamente utilizando `sort()`.

Después, se crea una lista llamada `counts` que contiene la cantidad de tweets correspondiente a cada etiqueta de `all_labels`. Se utiliza el diccionario `label_count` para obtener los conteos de cada etiqueta. Si una etiqueta no está presente en el diccionario, se asigna 0 como valor.

Seguidamente, se utiliza el objeto `ax3` para crear un gráfico de barras que muestra la cantidad de tweets por cada etiqueta. Se establecen las etiquetas en el eje x utilizando `all_labels` y se asignan los conteos en el eje y utilizando `counts`. Se establecen los ejes y títulos de las cantidades de tweets por etiqueta.

Después de configurar los gráficos, se utiliza `plt.tight_layout()` para ajustar la distribución de los subplots y evitar superposiciones. Luego, se muestra la gráfica utilizando `plt.show()`.

Por último, se guarda la figura generada en un archivo de imagen en la ubicación especificada utilizando `plt.savefig()`. En este caso, se guarda en la ruta "C:/Users/soler/Desktop/resultados.png".

6.2.7 Representación de resultados

Por último y muy relacionado con el punto anterior, una pequeña muestra de cómo se muestran los resultados finales. Para este caso hemos utilizado el usuario MarkHamill con un total de 10 tweets para una pequeña muestra.

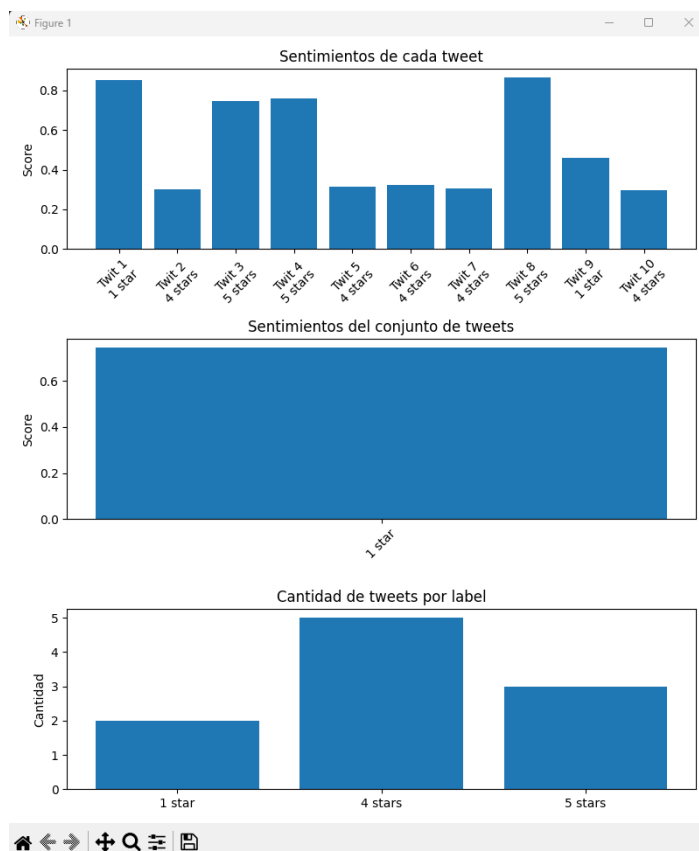


Gráfico 4 :Gráfico de resultados.

Fuente: Generación por programa.

```
=====
Tweet: One deserves to be elected again.

One deserves to be rejected again.

#ChooseWisely2024 https://t.co/qLGonZ6k3H
Label: 1 star
Score: 0.46194759011268616
```

Gráfico 5 :Muestra de análisis individual.

Fuente: Generación por programa.

Aún así, no son las únicas maneras de mostrar los resultados, el cliente puede solicitar cualquier otra que le convenga y se adapta el desarrollo a sus necesidades.

Una vez finalizada la explicación técnica del proyecto y su primera implementación de la solución, procedemos especificar como se ha desarrollado la elaboración del proyecto.

7. PLANIFICACIÓN DEL PROYECTO UTILIZANDO LA METODOLOGÍA SCRUM

Tal y como se ha visto en apartados anteriores, la metodología Scrum es un enfoque ágil para la gestión de proyectos que se centra en la flexibilidad, la colaboración y la entrega incremental de productos o resultados. En este apartado, se simulará la planificación seguida en el proyecto, teniendo como supuesto que todos los roles recaen sobre mi persona.

7.2 Definición del Producto y Objetivos

El producto que se va a desarrollar es una aplicación de análisis de sentimientos en Twitter. El objetivo principal es recopilar y analizar los últimos tweets de un usuario específico para determinar el sentimiento asociado a cada uno de ellos. La aplicación utilizará un modelo pre entrenado basado en Transformers para realizar el análisis de sentimientos.

Los objetivos específicos del proyecto son los siguientes:

- Recopilar y analizar los últimos tweets: La aplicación deberá acceder a la API de Twitter y obtener los últimos tweets del usuario especificado. Estos tweets serán utilizados como entrada para el análisis de sentimientos.
- Realizar el análisis de sentimientos: Se utilizará un modelo pre entrenado basado en Transformers para realizar el análisis de sentimientos en los tweets. El modelo clasificará cada tweet en una categoría de sentimiento, como positivo, negativo o neutro.

- **Mostrar los resultados del análisis:** Los resultados del análisis de sentimientos se mostrarán al usuario de manera clara y comprensible. Cada tweet será acompañado de su etiqueta de sentimiento correspondiente para facilitar la interpretación de los resultados.
- **Proporcionar evaluaciones individuales y globales:** Además de mostrar los resultados de manera individual para cada tweet, la aplicación también ofrecerá una evaluación global del conjunto de tweets analizados. Esto permitirá al usuario obtener una visión general de los sentimientos predominantes en los tweets.

7.3 Creación del Product Backlog

El Product Backlog es una lista priorizada de todos los elementos que representan los requisitos, funcionalidades y mejoras que se deben implementar en el proyecto. Estos elementos son definidos por el Product Owner en colaboración con el equipo de desarrollo y se ordenan en función de su importancia y valor para el cliente.

ID	Descripción	Prioridad	Esfuerzo
1	Obtener los últimos tweets de un usuario específico.	Alta	0,5 Sprint
2	Limpiar el texto de los tweets, eliminando URLs, menciones y hashtags no relevantes.	Media	0,5 Sprint
3	Implementar un clasificador de sentimientos utilizando Transformers y un modelo pre entrenado.	Alta	1 Sprint
4	Evaluar individualmente cada tweet y asignarle una etiqueta de sentimiento (positivo, negativo o neutro).	Alta	0,25 Sprint
5	Mostrar el resultado de la clasificación de cada tweet al usuario.	Alta	0,25 Sprint
6	Realizar una evaluación global del conjunto de tweets y proporcionar el sentimiento predominante.	Media	1 Sprint
7	Mejorar la eficiencia y el rendimiento del código de análisis de sentimientos.	Baja	0,25 Sprint
8	Implementar pruebas automatizadas para asegurar la calidad del código.	Alta	0,25 Sprint
9	Proporcionar documentación detallada sobre el uso del código y las instrucciones de configuración de las credenciales de Twitter.	Media	0,5 Sprint
10	Incorporar la posibilidad de realizar análisis de sentimientos en diferentes idiomas.	Media	0,5 Sprint
11	Implementar una interfaz de usuario intuitiva y amigable para facilitar la interacción del usuario con el código.	Opcional	1 Sprint
12	Mejorar la precisión del clasificador de sentimientos mediante ajustes y refinamientos del modelo pre entrenado.	Opcional	0,5 Sprint
13	Establecer alertas o notificaciones en caso de detectar tweets con sentimientos negativos importantes.	Opcional	0,25 Sprint
14	Realizar análisis de sentimientos en otros tipos de contenido, como comentarios de blog o reseñas de productos.	Opcional	0,5 Sprint

Cuadro 4: Product Backlog.

Fuente: Generación propia

Tomando como base este PB, se procede con el inicio de los denominados Sprint. En cada uno de ellos se realiza una lista de tareas acorde a cada elemento y se procede con su desarrollo, para finalmente realizar una reunión con la que informar y valorar el progreso realizado durante el Sprint. A continuación se desarrolla la planificación de tareas para cada uno de ellos.

7.4 Sprint 1

Sprint 1	
Tareas	Obtener los últimos tweets de un usuario específico
Investigar la API de Twitter	Realizar una investigación sobre la API de Twitter y comprender su funcionamiento.
Configurar las credenciales de autenticación de Twitter	Obtener y configurar las credenciales de autenticación necesarias para acceder a la API de Twitter.
Diseñar y desarrollar la lógica de obtención de tweets	Diseñar e implementar la lógica necesaria para obtener los últimos tweets de un usuario específico.
Manejar casos de error y excepciones	Implementar la gestión de casos de error y excepciones durante la obtención de tweets para garantizar la robustez.
Realizar pruebas unitarias	Ejecutar pruebas unitarias para verificar el correcto funcionamiento de la obtención de tweets.
Tareas	Limpiar el texto de los tweets, eliminando URLs, menciones y hashtags no relevantes
Eliminar URLs	Identificar y eliminar las URL presentes en el texto de los tweets utilizando expresiones regulares o funciones específicas de manipulación de texto.
Eliminar menciones de usuarios	Identificar y eliminar las menciones a otros usuarios en el texto de los tweets utilizando expresiones regulares o funciones específicas de manipulación de texto.
Eliminar hashtags no relevantes	Identificar y eliminar los hashtags no relevantes o que no aportan información significativa en el texto de los tweets utilizando expresiones regulares o funciones específicas de manipulación de texto.
Eliminar caracteres especiales y puntuación	Eliminar caracteres especiales y puntuación innecesarios en el texto de los tweets utilizando funciones específicas de manipulación de texto.

Convertir el texto a minúsculas	Convertir todo el texto de los tweets a minúsculas para una mayor coherencia y consistencia.
Eliminar espacios en blanco adicionales	Identificar y eliminar espacios en blanco adicionales o no deseados en el texto de los tweets utilizando funciones específicas de manipulación de texto.

Cuadro 5: Sprint 1.

Fuente: Generación propia

7.5 Sprint 2

Sprint 2	
Tareas	Implementar un clasificador de sentimientos utilizando Transformers y un modelo pre entrenado
Investigar modelos pre entrenados y bibliotecas de Transformers	Realizar una investigación sobre los modelos pre entrenados disponibles y las bibliotecas de Transformers compatibles.
Seleccionar un modelo pre entrenado adecuado	Evaluar y seleccionar un modelo pre entrenado que sea apropiado para el análisis de sentimientos en texto.
Configurar el entorno de desarrollo	Configurar el entorno de desarrollo con las bibliotecas y dependencias necesarias para trabajar con Transformers.
Diseñar e implementar la lógica del clasificador de sentimientos	Diseñar e implementar la lógica necesaria para utilizar el modelo pre entrenado y realizar el análisis de sentimientos.
Realizar pruebas unitarias y de integración	Ejecutar pruebas unitarias y de integración para verificar el correcto funcionamiento del clasificador de sentimientos.
Optimizar el rendimiento del clasificador	Realizar ajustes y optimizaciones para mejorar el rendimiento del clasificador de sentimientos.

Cuadro 6: Sprint 2.

Fuente: Generación propia

7.6 Sprint 3

Sprint 3	
Tarea	Evaluar individualmente cada tweet y asignarle una etiqueta de sentimiento
Preprocesamiento del texto del tweet	Realizar la limpieza del texto del tweet, eliminando URLs, menciones y hashtags no relevantes.
Utilizar el clasificador de sentimientos	Utilizar el clasificador de sentimientos previamente implementado para analizar el texto del tweet y asignarle una etiqueta de sentimiento.
Definir los criterios de asignación de etiquetas de sentimiento	Establecer los criterios y reglas para asignar las etiquetas de sentimiento (positivo, negativo o neutro) en función de los resultados.
Asignar la etiqueta de sentimiento al tweet	Aplicar los criterios definidos para asignar la etiqueta de sentimiento correspondiente a cada tweet evaluado.
Tarea	Mostrar el resultado de la clasificación de cada tweet al usuario
Diseñar la interfaz de usuario	Diseñar la interfaz de usuario donde se mostrarán los resultados de la clasificación de cada tweet.
Obtener los resultados de clasificación	Obtener los resultados de la clasificación de cada tweet realizado previamente.
Formatear los resultados de clasificación	Formatear los resultados de clasificación para que sean legibles y comprensibles para el usuario.

Mostrar los resultados de clasificación en la interfaz de usuario	Mostrar los resultados de clasificación de cada tweet en la interfaz de usuario de manera clara y estructurada.
Proporcionar una evaluación individual de cada tweet	Presentar la etiqueta de sentimiento asignada a cada tweet, junto con información adicional si es necesario (por ejemplo, puntaje).
Tarea	Realizar una evaluación global del conjunto de tweets y proporcionar el sentimiento predominante
Recopilar los resultados de clasificación de los tweets	Obtener los resultados de clasificación de sentimiento para cada tweet del conjunto evaluado.
Analizar los resultados de clasificación	Analizar los resultados de clasificación para identificar los diferentes sentimientos asignados a los tweets (positivo, negativo, neutro).
Calcular la frecuencia de cada sentimiento	Contabilizar la frecuencia de cada sentimiento asignado a los tweets (positivo, negativo, neutro) para obtener una visión global.
Determinar el sentimiento predominante	Identificar el sentimiento que tiene una frecuencia mayor en el conjunto de tweets evaluado.
Proporcionar el sentimiento predominante al usuario	Mostrar al usuario el sentimiento predominante del conjunto de tweets y brindar una breve descripción o resumen.

Cuadro 7: Sprint 3.

Fuente: Generación propia

7.7 Sprint 4

Sprint 4	
Tarea	Mejorar la eficiencia y el rendimiento del código de análisis de sentimientos
Realizar un análisis de rendimiento	Evaluar el rendimiento actual del código de análisis de sentimientos para identificar posibles cuellos de botella o áreas de mejora.
Optimizar el procesamiento del texto de los tweets	Identificar oportunidades de optimización en el preprocesamiento del texto de los tweets, como utilizar técnicas más eficientes de limpieza.
Implementar técnicas de procesamiento paralelo o distribuido	Explorar la posibilidad de utilizar técnicas de procesamiento paralelo o distribuido para acelerar el análisis de sentimientos en grandes volúmenes de datos.
Optimizar la carga y procesamiento de modelos pre entrenados	Investigar y aplicar técnicas de optimización para reducir el tiempo de carga y procesamiento de los modelos pre entrenados utilizados en el análisis de sentimientos.
Implementar almacenamiento en caché de resultados intermedios	Utilizar técnicas de almacenamiento en caché para evitar el procesamiento repetitivo de los mismos datos y acelerar el tiempo de respuesta.
Realizar pruebas de rendimiento y ajustes iterativos	Realizar pruebas exhaustivas de rendimiento en diferentes condiciones y ajustar el código según los resultados para mejorar continuamente la eficiencia y el rendimiento.
Tarea	Implementar pruebas automatizadas para asegurar la calidad del código
Identificar los componentes y funcionalidades clave	Identificar los componentes y funcionalidades críticas del código de análisis de sentimientos que requieren pruebas automatizadas.

Diseñar casos de prueba	Definir los casos de prueba que cubran los diferentes escenarios y funcionalidades del código.
Implementar las pruebas unitarias	Desarrollar pruebas unitarias para verificar el funcionamiento individual de los componentes del código.
Implementar las pruebas de integración	Crear pruebas de integración para verificar la correcta interacción entre los diferentes componentes del sistema.
Configurar pruebas de regresión	Establecer pruebas de regresión para asegurar que las modificaciones o mejoras realizadas no introduzcan errores en funcionalidades existentes.
Automatizar la ejecución de las pruebas	Implementar un marco o herramienta de pruebas automatizadas que permita ejecutar las pruebas de manera eficiente y repetible.
Monitorear la cobertura de las pruebas	Evaluar la cobertura de las pruebas y garantizar que se hayan abarcado todas las funcionalidades críticas del código.
Ejecutar las pruebas de forma continua	Configurar una integración continua para ejecutar las pruebas automáticamente después de cada cambio o actualización en el código.
Analizar y corregir los fallos identificados durante las pruebas	Analizar los resultados de las pruebas, identificar los fallos y corregirlos de manera oportuna para asegurar la calidad del código.
Tarea	Proporcionar documentación detallada sobre el uso del código y las instrucciones de configuración de las credenciales de Twitter
Documentar el proceso de configuración de las credenciales	Detallar los pasos necesarios para obtener las credenciales de autenticación de Twitter, incluyendo la creación de una aplicación en la plataforma.

Explicar cómo ingresar las credenciales en el código	Indicar en qué parte del código se deben ingresar las credenciales obtenidas y cómo se deben estructurar para que la autenticación sea exitosa.
Describir los requisitos del entorno de ejecución	Enumerar los requisitos del entorno en el que se ejecutará el código, como las versiones de Python y las librerías externas necesarias.
Detallar el proceso de instalación y configuración	Explicar los pasos necesarios para instalar las dependencias del proyecto y configurar el entorno para ejecutar el código de manera correcta.
Describir los parámetros y opciones de personalización	Documentar los parámetros y opciones disponibles para personalizar el comportamiento del código, como el número máximo de tweets a analizar, por ejemplo.
Proporcionar ejemplos de uso	Incluir ejemplos concretos de cómo utilizar el código en diferentes escenarios, mostrando los comandos y argumentos necesarios para obtener resultados.
Explicar el formato de los resultados	Describir el formato en el que se presentarán los resultados del análisis de sentimientos, como etiquetas de sentimiento y puntuaciones asociadas.
Tarea	Incorporar la posibilidad de realizar análisis de sentimientos en diferentes idiomas
Investigar modelos de análisis de sentimientos	Realizar una investigación sobre modelos pre entrenados y librerías disponibles que admitan el análisis de sentimientos en diferentes idiomas.
Seleccionar un modelo adecuado	Evaluar los modelos disponibles y seleccionar aquellos que sean adecuados para realizar el análisis de sentimientos en diferentes idiomas.

Configurar el clasificador para idiomas específicos	Realizar las configuraciones necesarias en el clasificador para admitir diferentes idiomas, como cargar modelos específicos para cada idioma.
Realizar pruebas con muestras de diferentes idiomas	Obtener muestras de texto en diferentes idiomas y realizar pruebas para asegurar que el análisis de sentimientos funciona correctamente.
Ajustar los resultados según el idioma	Realizar ajustes en los resultados del análisis de sentimientos según las particularidades lingüísticas y culturales de cada idioma.
Actualizar la documentación	Modificar la documentación para incluir instrucciones sobre cómo utilizar el análisis de sentimientos en diferentes idiomas.

Cuadro 8: Sprint 4.

Fuente: Generación propia

7.8 Sprint futuros

Una vez realizada la funcionalidad básica del proyecto, se dedicarán los esfuerzos de futuros Sprint a ir puliendo el mismo e incluyendo las funcionalidades opcionales registradas como tal en el Product Backlog para aportar valor adicional al software.

Estas funcionalidades adicionales junto algunos fallos detectados van a ser desarrollados como puntos de mejora a implementar en el futuro que se van a desarrollar en el apartado siguiente.

8.POSIBLES MEJORES DE LA IMPLEMENTACIÓN

En toda implementación software una vez se finaliza su primer desarrollo quedan aspectos en el tintero para desarrollar o mejorar. Evidentemente, este proyecto no es la excepción, pues existen miles de posibilidades a mejorar.

Partiendo de la base de que está pensado en su origen como una herramienta de muestra y como base para futuras implementaciones de proyectos con necesidades específicas, se trata de una versión con poca interacción de un usuario, por lo que la prioridad de las mejoras está enfocada a mejorar este ámbito.

8.1 Parametrización

Tal y como hemos comentado, es necesaria más interacción usuario-máquina, por lo que el primer paso es desarrollar código que permita introducir al usuario cada vez los parámetros con los que desea trabajar su análisis, y no acceder a ellos a nivel de código para después volver a compilar y ejecutar.

Una sucesión de preguntas al usuario durante la ejecución en cada paso para que sea él quién parametriza cada apartado sería el enfoque adecuado. Por ejemplo, permitiendo la personalización de las gráficas o de las etiquetas del modelo.

8.2 Integración de otras fuentes de datos

Además de utilizar la API de Twitter para obtener los tweets, convendría expandir la funcionalidad del código para integrar otras fuentes de datos. Esto permitiría realizar análisis de sentimientos en textos provenientes de diferentes plataformas y contextos. Algunas otras fuentes de datos a incluir:

Integración de Facebook: Utilizar la API de Facebook para obtener publicaciones, comentarios o mensajes de usuarios y aplícales análisis de sentimientos. Considerar diferentes tipos de contenido, como publicaciones de usuarios, comentarios en páginas, mensajes privados, etc.

Integración de Reddit: Utilizar la API de Reddit para acceder a subreddits relevantes y extraer publicaciones y comentarios. Realizar el análisis de sentimientos en estos textos para obtener una comprensión de la opinión de los usuarios en diferentes comunidades de Reddit.

Integración de noticias o blogs: Configurar un sistema de web scraping para obtener noticias o artículos de blogs relacionados con un tema específico. Aplicar el análisis de sentimientos a estos textos para tener una visión general de la percepción pública en relación con ese tema.

Integración de encuestas en línea: Utilizar APIs de plataformas de encuestas en línea para recopilar respuestas de usuarios. Estas respuestas pueden estar relacionadas con productos, servicios o temas específicos. Realizar el análisis de sentimientos en los comentarios de las encuestas para obtener información sobre las opiniones y actitudes de los participantes.

Al ampliar la integración de otras fuentes de datos, es importante para el proyecto tener en cuenta las limitaciones y condiciones específicas de cada plataforma o API. Al igual que con Twitter, es necesario cumplir con las políticas y términos de uso de cada fuente de datos y de obtener los permisos necesarios para acceder y utilizar la información de manera adecuada y ética.

8.3 Inclusión de Bases de Datos

Otro añadido de gran valor para el software sería la capacidad de poder trabajar también con bases de datos e integrarse con las mismas, para no sólo trabajar con muestras recogidas en el momento de la ejecución.

Como parte de la puesta en marcha de esta mejora, podría considerarse la migración a un entorno más amigable con estos sistemas, como puede ser Java.

8.4 Fine-tuning del Modelo

Una posible mejora para el proyecto es realizar el fine-tuning del modelo pre entrenado utilizado. Esto implica recopilar datos etiquetados específicos del dominio o idioma de interés y utilizarlos para ajustar los parámetros del modelo. El conjunto de datos debe incluir ejemplos de sentimientos positivos, negativos y neutros para un entrenamiento equilibrado. Una vez ajustado el modelo, se evalúa su desempeño utilizando un conjunto de prueba y se pueden realizar ajustes adicionales mediante validación cruzada. Finalmente, se integra el modelo fine-tuned en la aplicación de análisis de sentimientos, reemplazando el modelo pre entrenado original. Este proceso de fine-tuning permite adaptar el modelo a las características específicas del dominio o idioma, mejorando así su precisión y capacidad de generalización en el análisis de sentimientos.

Es importante tener en cuenta que el éxito del fine-tuning depende de la disponibilidad de datos etiquetados relevantes y suficientes, así como de los recursos necesarios para el entrenamiento. Sin embargo, esta mejora puede ayudar a optimizar el rendimiento del modelo y obtener resultados más precisos en el análisis de sentimientos.

Una ventaja del fine-tuning en comparación con el entrenamiento completo de un modelo desde cero es la reducción del tiempo requerido. El proceso de entrenamiento completo puede llevar mucho tiempo y requiere una gran cantidad de recursos computacionales. En cambio, el fine-tuning se basa en un modelo pre entrenado existente y solo se ajustan los parámetros específicos del dominio o idioma de interés. Esto implica un tiempo de entrenamiento considerablemente menor, lo que permite una implementación más rápida del modelo de análisis de sentimientos en Twitter. Además, el fine-tuning aprovecha el conocimiento previo del modelo pre entrenado, lo que puede resultar en un rendimiento inicialmente mejorado en comparación con un modelo entrenado desde cero.

8.5 Interfaz

Sin duda alguna, la mayor característica a echar en falta de la herramienta es una interfaz que permita su uso de manera amigable para cualquier usuario.

Una buena aproximación a lo que sería ideal, es una pantalla de inicio dónde el software te da la bienvenida y te solicita los primeros parámetros de trabajo. El usuario especifica si desea buscar por nickname o por hashtag para a continuación especificar el criterio de búsqueda y la cantidad.

Una vez comienza el proceso, el propio programa va guiando sobre los resultados que se van obteniendo, y vuelve a solicitar cómo se desea la información, si en un tipo de gráfico u otro, además de especificar las características del mismo en cuanto ejes, nomenclatura etcétera.

A continuación, se presentan algunas maquetas de ideas para la interfaz:

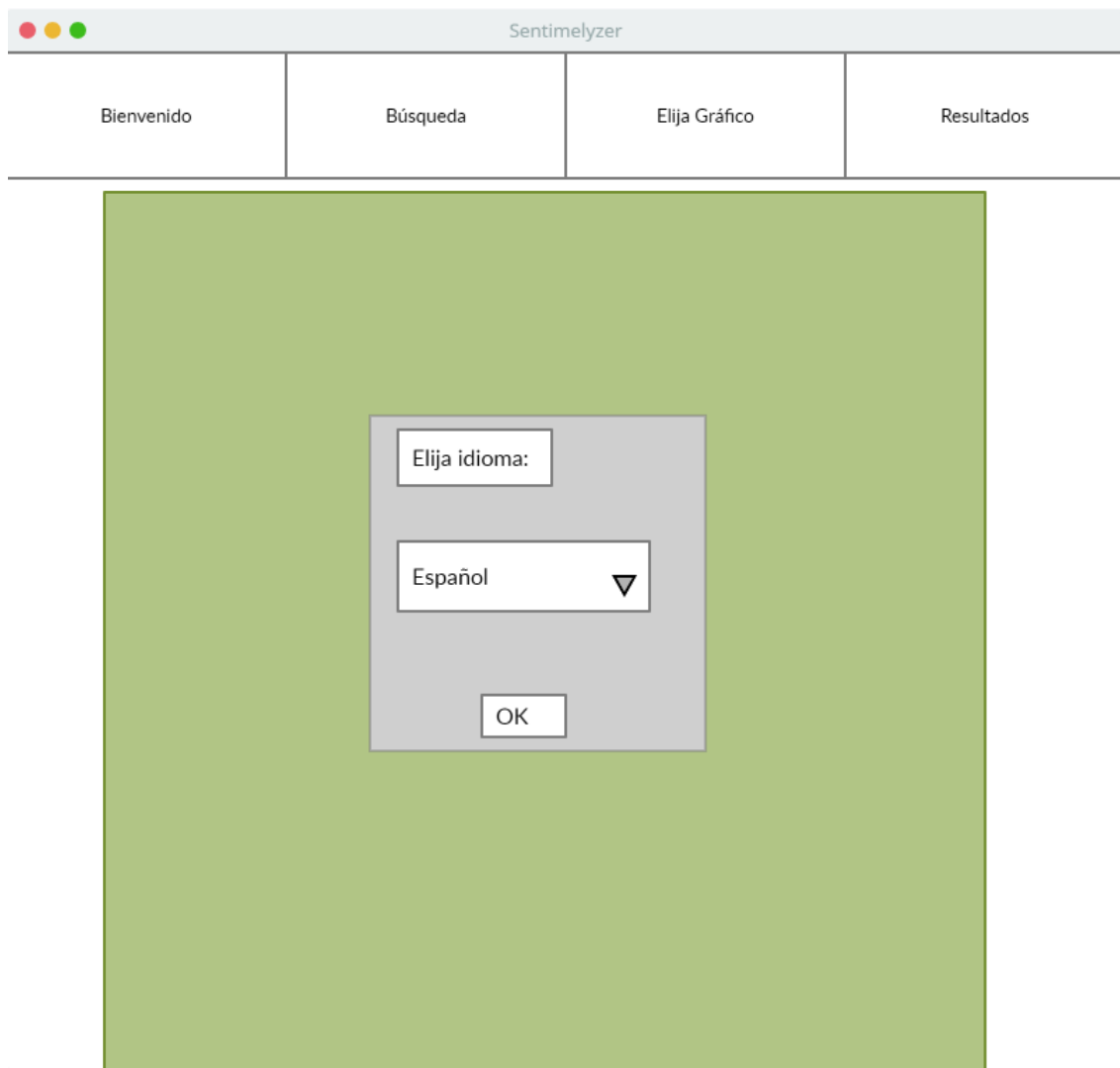


Imagen 16 :Interfaz de bienvenida.

Fuente: Elaboración propia

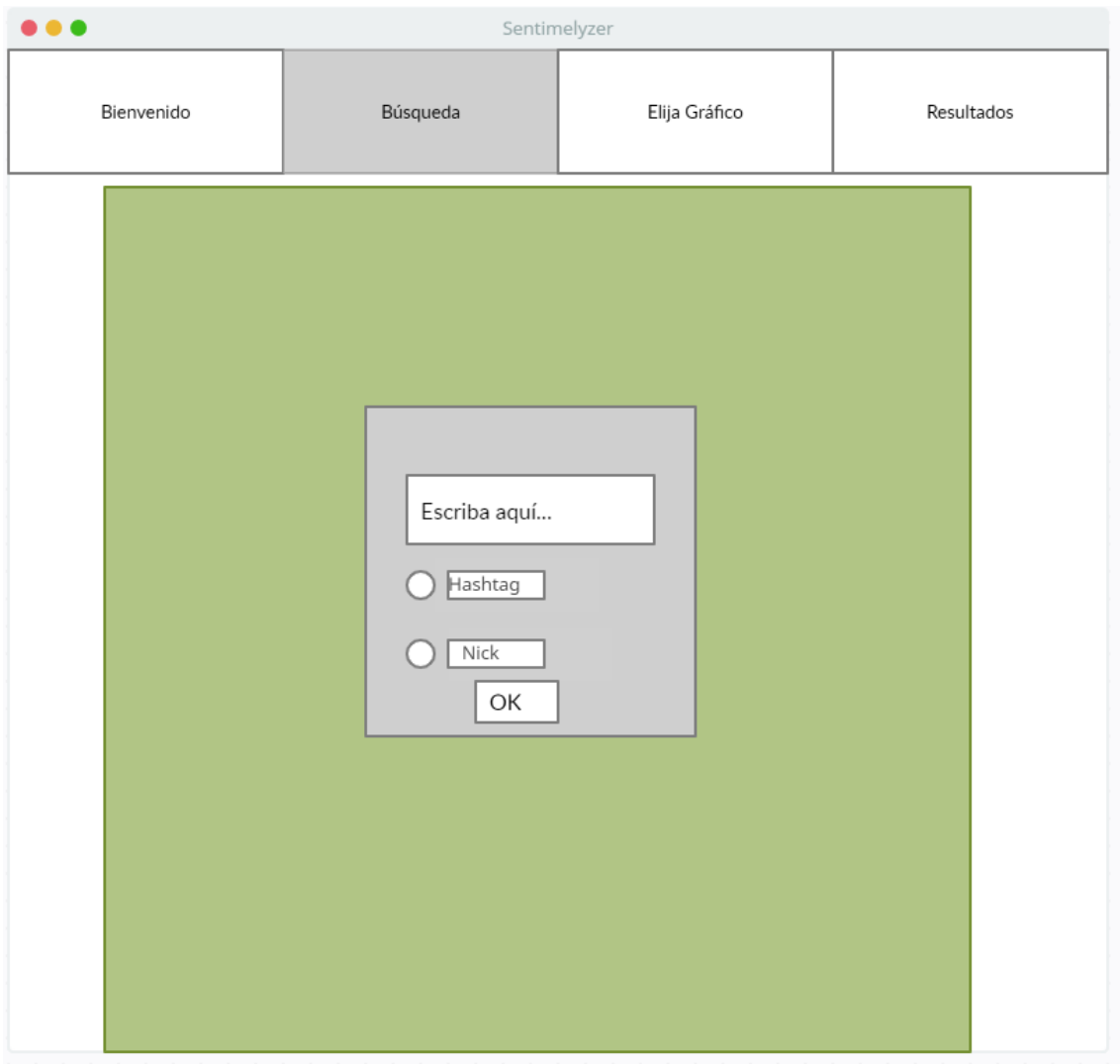


Imagen 17 :Interfaz de búsqueda.

Fuente: Elaboración propia

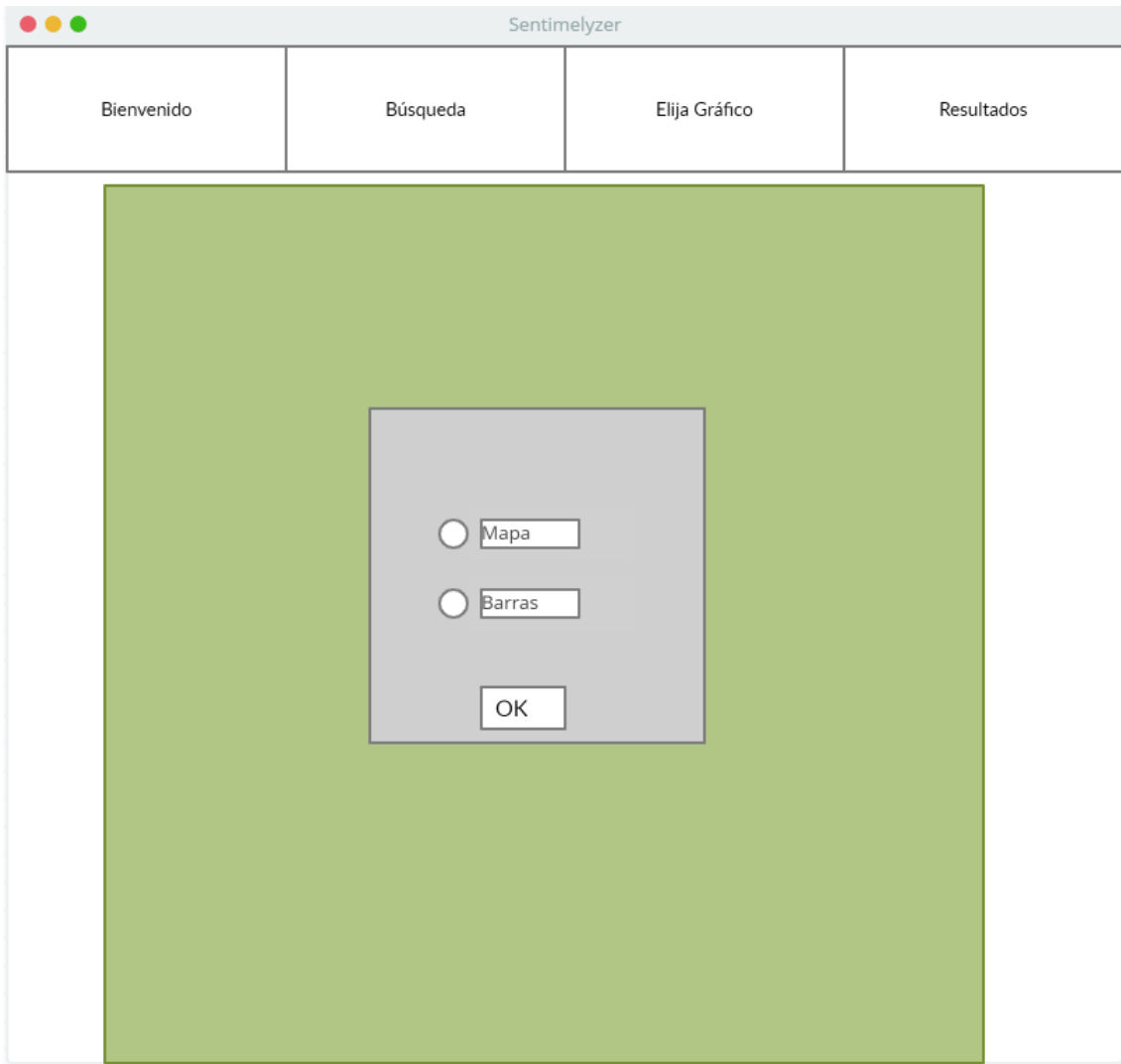


Imagen 18 :Interfaz de selección.

Fuente: Elaboración propia

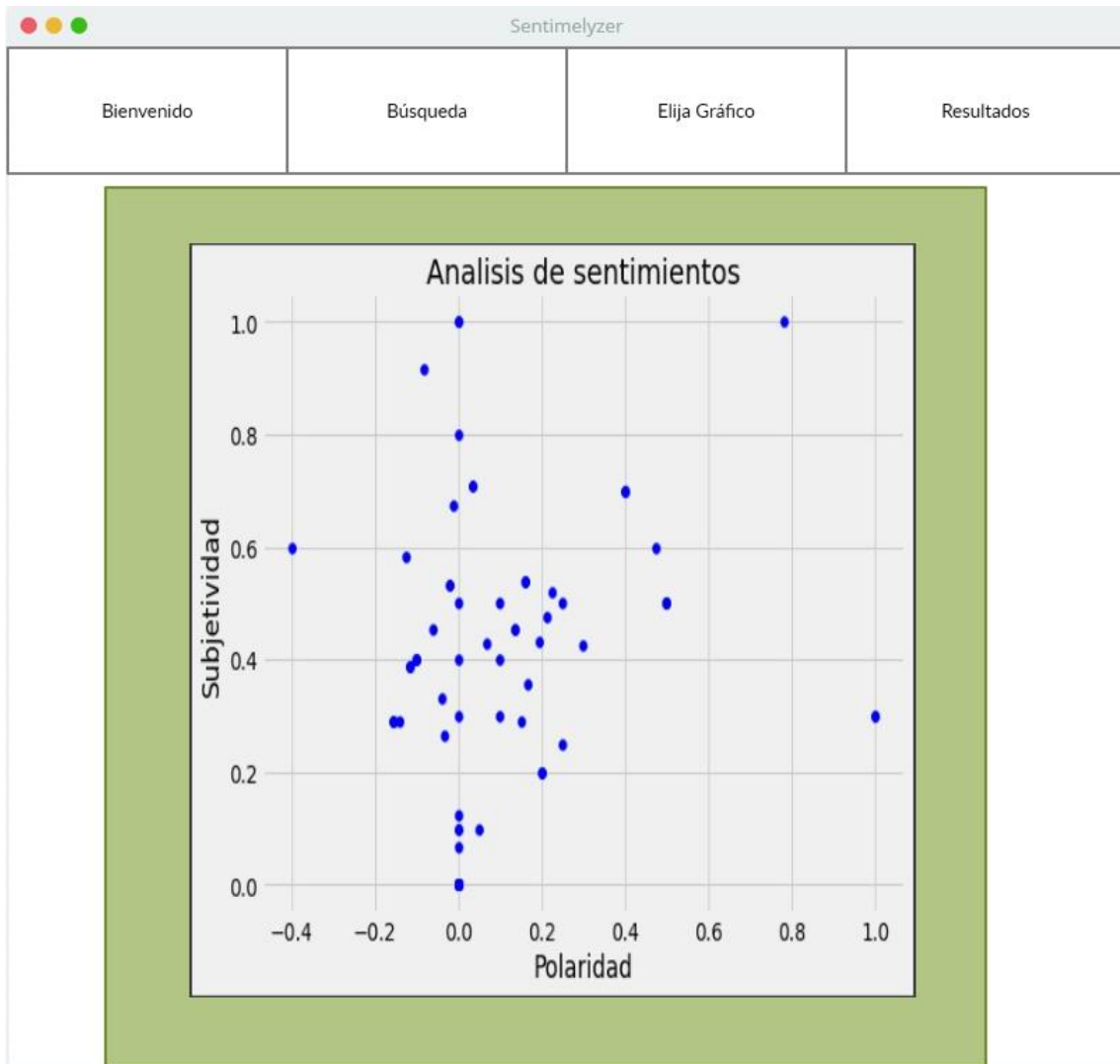


Imagen 19 :Resultados.

Fuente: Elaboración propia

8.6 Corrección de errores

Como último apartado de mejoras, es el turno de la corrección de errores, tanto los detectados por la prueba como los posibles generados por las nuevas implementaciones.

Implementar manejo de excepciones para capturar posibles errores durante la interacción con la API de Twitter y el análisis de sentimientos. Esto permitirá gestionar de manera adecuada situaciones inesperadas, como falta de conexión a Internet o límites de solicitud.

Finalizada las anotaciones en cuanto a posibles mejoras de la implementación, llegan las partes finales del trabajo donde se especifica la estrategia a seguir para el proyecto.

9. ÁMBITO ÉTICO-LEGAL

Un aspecto que no podemos olvidar a la hora de realizar un desarrollo software es la legalidad vigente en el área que trabaja. Esta legalidad es la que nos va a marcar los límites de lo que podemos o no podemos hacer, influenciando así en los requisitos y funcionalidades de toda aplicación.

Además, en particular esta área, puesto que vamos a trabajar con redes sociales y sus usuarios, tenemos un tema conflictivo, pues se trabaja con información bastante sensible. Por suerte, en Europa estamos bastante avanzados en el aspecto legal de Internet en comparación a otras partes del mundo, y aunque aún queda mucho camino por recorrer, tenemos en España la Ley Orgánica de Protección de Datos de Carácter Personal (LOPDGDD).

La finalidad de la LOPD es proteger al individuo y su intimidad y privacidad. Para ello legisla el uso y tratamiento de datos de carácter personal, las libertades públicas y los derechos fundamentales de las personas. El organismo encargado por el cumplimiento de esta Ley es la Agencia Española de Protección de Datos.

Para poder conocer cómo nos afecta a nosotros en cuanto a nuestro proyecto software, debemos recurrir a la política de Privacidad de Twitter.

9.1 Política de privacidad de Twitter

Si acudimos a la página oficial que describe las políticas de Twitter, podemos acceder al apartado de datos compartidos. Según la propia Twitter, categoriza en 5 formas generales la compartición de datos:

- 1. Cuando el usuario Twitteo, publica contenido y comparte:** En este caso, la información es de carácter público, contenido compartido e información básica del perfil tal como nombre de usuario, pseudónimo y foto de perfil.

2. Cuando lo exige la ley, para evitar daños o cuando es de interés público:

Twitter puede compartir o divulgar datos de usuario si es razonable para:

- a. Cumplimiento de leyes, normativas o requerimientos judiciales**
- b. Velar por la seguridad de cualquier persona, así como de la integridad de la plataforma**
- c. Explicar el borrado de contenidos o cuentas de sus servicios**
- d. Abordar problemas técnicos o de fraude y seguridad**
- e. Proteger sus derechos o bienes y de aquellos que usan sus servicios**

3. Con sus entidades vinculadas: Pueden compartir datos con sus entidades vinculadas para ofrecer productos y servicios

4. Como consecuencia de un cambio de titularidad: Pueden compartir, vender o transferir datos de los usuarios en el contexto de una fusión, adquisición, reorganización o venta de activos y procesos similares.

5. Con terceros o integraciones de terceros: Twitter comparte datos con proveedores de servicios y anunciantes como forma de negocio. En cuanto a contenido de terceros e integraciones, comparten datos según las preferencias marcadas por el usuario en su cuenta.

Por último, el tema que más nos concierne, en cuanto a API's e integraciones, Twitter permite todos aquellos contenidos que sean públicos. Citando directamente de la fuente: “Utilizamos tecnologías como [API](#) e integraciones para permitir que los contenidos públicos de Twitter estén disponibles para su utilización por sitios web, aplicaciones y otros terceros, por ejemplo, mostrando Tweets en un sitio web de noticias o analizando lo que la gente dice en Twitter. Por lo general, permitimos la utilización de estos contenidos en cantidades limitadas de forma gratuita y cobramos cánones de licencia si el acceso se realiza a gran escala. Disponemos de [condiciones generales](#) que regulan la utilización de esta información y un programa de cumplimiento normativo para velar por el cumplimiento de estas condiciones. No obstante lo anterior, dado que estas personas y empresas no están vinculadas a Twitter, puede que los contenidos que muestren no reflejen las actualizaciones que hayas introducido en Twitter.”

De entre las condiciones generales, uno de los motivos es el trabajo académico y el ámbito de negocio, por lo que no tenemos problemas con Twitter y utilizar los tweets públicos del Time Line para nuestro uso.

9.2 Aspectos éticos y morales

Ya hemos comprobado que no existe problema legal en el uso de la información que utilizamos para la aplicación. Sin embargo, otra de las preguntas más importantes que debe hacerse todo desarrollador es: ¿Es ético lo que estoy haciendo?

Todo el mundo tiene una opinión acerca de cómo quiere que traten su privacidad. Existen personas que prefieren mantener la confidencialidad al máximo posible y existen personas que son totalmente opuestas, no solo no les importa sino que en muchos casos les gusta compartir. Si bien es cierto que parte del debate se basa en que cuando una persona se convierte en usuario de redes sociales, acepta sus términos y servicios y es por propia voluntad el que comparte con los demás; en la mayoría de casos no están alineadas la intencionalidad y los motivos por los que un usuario comparte una información y el tratamiento que recibe después por parte de compañías o terceros.

Además, entra en juego un factor que hace más personal e íntima la información, y es que en nuestro software tratamos de encontrar sentimientos del emisor en su publicación. En general todavía no se han superado los tabúes y estigmas de hablar acerca de nuestros sentimientos en la sociedad actual, por lo que muchas personas considerarían esto un ataque directo a su intimidad, pues lo consideran un tema muy interior. En parte, sí es un poco inmoral si se utiliza con intención de negocio, pues uno de los temas candentes es el límite de la publicidad y la intrusión de los negocios en nuestras vidas, y si tratamos de ponerles límites en acciones mucho menos intrusivas, es de esperar que no guste si se trata de sentimientos.

Ahora bien, ¿y si actuamos por una causa social? ¿Y si no es el dinero la motivación de la aplicación?

Cada vez hay más conciencia de la importancia de la salud de mental, y muchas personas están de acuerdo en que no se valora lo suficiente y no nos educan nada en inteligencia emocional. Las redes están llenas de personas que les cuesta mostrar sus sentimientos, en especial los negativos. Mucha gente necesita ayuda y no puede o no sabe pedirla, y el resto de usuarios tampoco somos capaces de detectar estos problemas. Y ocurren delante de nosotros, en nuestras pantallas muy a menudo, uno detrás de otro.

Por esto el desarrollo de sistemas de detección y prevención serían una buena incorporación al conjunto de herramientas que nos ofrece la tecnología. A nadie le gustaba tener que conducir con cinturón porque era incómodo, pero se acabó aceptando por el bien de la seguridad vial. Si añadimos tantos mecanismos de seguridad física en todos los ámbitos (riesgos laborales, seguros, sistemas de protección vial, alarmas de seguridad...) ¿por qué no añadir también mecanismos de seguridad psicológica y emocional?

Como conclusión de este apartado y opinión personal, este sería uno de los debates primordiales en cuanto a materia de redes sociales ahora que ya tenemos camino recorrido en cuanto a datos personales. Si hemos trabajado duro y conseguimos avances por la protección de la información de las personas en Internet, el siguiente paso es proteger algo más valioso aún, a las personas mismas.

10.IMPLEMENTACIÓN DE LA ESTRATEGIA

Ha llegado el momento de implantar la estrategia escogida tras los análisis realizados anteriormente, detallando las acciones que se van a llevar a cabo para ello. Aunque la estrategia sea buena y sea la elección correcta, si no se implanta de manera adecuada puede acabar no teniendo el resultado esperado.

10.1 Estrategia seleccionada

Dicha estrategia de segmentación se fundamenta en la diferenciación respecto a los competidores del sector, definida tal que:

Ofrecer un servicio personalizado de desarrollo de software basado en el análisis de sentimientos con un proceso de elaboración personal y directo desde el primero momento, cada línea de código escrita atendiendo las necesidades específicas del cliente.

El objetivo es transmitir valor añadido al cliente ofreciéndole un software completamente ideado para él, nada de software genéricos con una capa de personalización por encima para que el cliente lo sienta suyo. Durante todo momento mantener el contacto continuo y estar dispuesto a ofrecer el mejor servicio y en definitiva, que el cliente se sienta escuchado y con las riendas del desarrollo de su software.

Con la estrategia definida, es el momento de concretar las acciones a tomar para llevarla a cabo.

10.2 Plan de implantación

Ahora que tenemos finalizados todos los análisis necesarios y hemos definido todas las estrategias a seguir en función a ellos y sus resultados, es el momento de definir el plan de implantación con las actividades que van a llevarse a cabo.

10.2.1 Formación formal de la sociedad

En primer lugar, para poder empezar de manera formal con el proyecto, es necesario dar el primer paso y formar una sociedad e inscribirla en el registro, creando una marca propia a la que asociarla.

Una vez logrado, el siguiente paso es la creación de una página web presentación con la que mostrar la información básica del proyecto, añadir la información de contacto y servir como fuente de redirección de la campaña publicitaria.

10.2.2 Búsqueda de financiación

Tal y como se había comentado en puntos anteriores, la principal fuente de financiación inicial va a residir en solicitar a la administración pública una subvención enfocada en el desarrollo de PYMES y desarrollo tecnológico.

Dichas subvenciones se solicitan directamente desde la página web de la GVA.

10.2.3 Publicidad en redes sociales

Contando con que los recursos son limitados la opción con menos desembolso económico son las redes sociales.

En primer lugar, recurrimos a las redes sociales basadas en formato de texto más utilizadas como son Twitter y Facebook. Para ellas, el plan de márketing ideado reside en preparar junto a un psicólogo profesional una batería de mensajes virales en dichas redes y un análisis de la intención de los mismos. Mensajes que a simple vista no son trascendentes pero que un análisis por parte de un profesional sea capaz de revelar intencionalidad emocional variada, desde emoción, alegría y felicidad hasta emociones más negativas como tristeza, melancolía o síntomas de depresión.

Por otra parte, desde el apartado técnico, promocionaremos el código de la prueba técnica de nuestro software en GitHub, pues la plataforma aparte de servir de entorno de trabajo para desarrolladores, también hace a su vez de punto de encuentro de opiniones del sector y funciona como una pequeña red social. Con esto se busca atraer desarrolladores que les interese el proyecto para aportar su opinión o considerar una incorporación al equipo.

10.2.4 Contacto directo con el nicho objetivo

Adicionalmente a la campaña de publicidad, para tratar de maximizar las posibilidades de éxito en el inicio del proyecto, una de las acciones clave es contactar directamente con el nicho definido como objetivo.

La administración pública está abierta a recibir contacto en todo momento, por lo que acudir a la sede de la GVA en Valencia y concertar una cita donde explicar el proyecto con la demostración técnica es una posibilidad real.

10.2.5 Captación de talento y formación de un equipo

La situación inicial del proyecto cuenta con una sola persona, y cabe esperar que si van teniendo éxito el resto de acciones, haya demanda de servicios y sea necesario ampliar el equipo. Como no hay recursos como para contratar personal, la intención es buscar perfiles similares a los del gerente del proyecto, alumnos recién egresados con conocimientos listos para poner en práctica y ganas de emprender, y ofrecerles así formar parte en forma de socios.

Finalizado este desarrollo de plan de implementación a seguir, y habiendo analizado previamente la estrategia, es el momento de empezar a actuar y dar comienzo a todo el trabajo realizado.

11.RESULTADOS INICIALES DEL PROYECTO

Con la puesta en marcha del proyecto es momento de analizar los primeros resultados obtenidos de toda la estrategia planteada.

En primer lugar, el software de la prueba técnica está completado con sus funcionalidades básicas y transmite la idea del proyecto y de lo que se puede alcanzar a hacer. Además está listo para subirse a GitLab y comenzar a darle repercusión. El nombre de la aplicación y de la marca se ha definido como “Sentimelyzer” para darle registro.

En cuanto a la financiación, han sido recogidos los primeros documentos de información acerca de cómo solicitar la subvención correspondiente. De paso, ha sido solicitada una entrevista en Valencia para la presentación de la idea, a esperas de obtener respuesta.

Por otro lado, referido a la campaña publicitaria, han sido contactado diversos profesionales de la psicología y están elaborando presupuestos de sus servicios.

Por último, en cuanto al equipo, han sido contactados 6 alumnos recién egresados y 2 de ellos están interesados en el proyecto, pasarán a formar parte del proyecto en forma de socios y está definir las aportaciones de cada uno.

11.1 Sigüientes pasos

Una vez tomados los primeros resultados, todavía quedan por cumplir el resto de las actividades planteadas en la estrategia y su plan de implantación.

Las prioritarias son la solicitud de la subvención para tener financiación con la que dar comienzo a la publicidad en primer lugar. En segundo plano, la confección de la página web y reunirse con los nuevos socios.

En tercer lugar, llevar a cabo todas las mejoras que se puedan lograr del apartado de mejoras a la implementación software.

Por último, gran parte del futuro del proyecto depende de la respuesta recibida por parte de la Generalitat y dependerá de si se sigue adelante o no.

12.CONCLUSIONES

Con la elaboración de este trabajo he experimentado de primera mano toda la problemática que afronta una persona con una idea a la que quiere dar forma y busca darle salida y emprender con ella. En definitiva, ha sido la experiencia que viven cientos de PYMES en España en estos tiempos marcados por crisis consecutivas, incertidumbre y miedo.

El objetivo del proyecto era profundizar en un área de conocimiento que me interesa enormemente como es el análisis de sentimiento y el *machine learning* y desarrollar una estrategia que fuera capaz de dar inicio a un proyecto de desarrollo de software en el ámbito. Para ello he puesto en uso multitud de los conocimientos adquiridos durante estos años de estudio.

En primer lugar, me he documentado acerca del tema en cuestión para formar una base teórica sólida donde trabajar un análisis tanto del entorno como del proyecto y la idea para determinar qué factores podrían ser determinantes en el éxito o fracaso del mismo.

Con la información obtenida del análisis, he sido consciente de lo duro y competitivo que es el sector y cómo es realmente difícil y complejo iniciarse de manera correcta en el mundo. Está lleno de impedimentos y si partes de una base de recursos limitados tienes más en tu contra que a tu favor y debes explotar al máximo cada pequeña oportunidad.

En base a ese análisis, he desarrollado la estrategia que consideré más factible dados los resultados del mismo y la serie de tareas a realizar para su implementación. Sin embargo, en esta estrategia reside una buena cantidad de optimismo. Siendo objetivos, se trata de un proyecto ambicioso para buscar obtener rentabilidad económica del mismo, pues para poder lograrlo debería emprenderse con muchos más recursos que una idea y una persona con voluntad. A esto hay que añadirle que es un área de conocimiento compleja que requiere de años de estudio para manejarse con soltura.

Por tanto, llegado este punto concluyo que es un proyecto destinado o bien a quedarse como una afición académica y un experimento, o bien como una idea a intentar pero sin arriesgar mucho. Tal y como se ha visto durante el desarrollo del mismo, la prudencia es la base de actuación, si no se recibiera una buena propuesta de desarrollo a manos de la administración pública o de un cliente privado con una mentalidad similar, a realizar un proyecto social que no le importe obtener rentabilidades económicas convendría directamente no seguir con el mismo y cancelar su progreso. Esto se debe en gran medida, a que para poder arrancar requiere de comprometerse en muchas tareas distintas tanto en tiempo como en inversión de unos recursos que no se disponen, por lo que hay más posibilidades de fracaso y endeudamiento que de éxito.

Sin embargo, no descarta que haya sido una experiencia bonita y llena de aprendizaje, la satisfacción de intentar dar forma a tus ideas y aspiraciones es muy gratificante y pocas cosas consiguen causar en ti ese sentimiento tan particular, en especial cuando estás utilizando nuevos métodos y nuevas herramientas a las ya conocidas simplemente por el hecho de tener curiosidad y querer aprender más sobre nuevos elementos. Como añadido, recalcar que el hecho de que sea poco viable hoy día, no significa que sea una renuncia total a la idea, si en algún momento del futuro cuento con más recursos económicos u oportunidades de inversión o financiación puede replantearse el proceso de análisis y desarrollo de una estrategia nueva y ver si es más factible su desarrollo.

Por último y no menos importante, aprovecho para remarcar de nuevo la importancia de la salud mental en estos días. Se trata de un tema que ha pasado por muchos tabúes y se deja en segundo plano en la sociedad actual, y es tan importante como la salud física. Hoy día existen infinidad de elementos que atacan a nuestra salud mental y generan daños en ella, que se acaban manifestando como enfermedades que pueden tener graves consecuencias y pasan muy desapercibidas. Todos debemos aportar a la concienciación de esta realidad y promover acciones que ayuden al cuidado de los demás, como el fomento de un uso responsable de las redes sociales tanto por parte de los usuarios como de las compañías que las dirigen.

13.BIBLIOGRAFÍA

Gráfico del precio de la energía obtenido en www.statista.com

Gráfico del desempleo obtenido en www.statista.com

Gráfico del acceso a teléfonos móviles obtenido en www.statista.com

Uso de las redes sociales en España, <https://unavidaonline.com/estadisticas-redes-sociales/>

Competidores del sector del análisis de sentimiento,
<https://www.capterra.es/directory/30933/text-mining/software>

Información de subvenciones GVA,
https://www.gva.es/es/web/temesgva/inicio/empresas/emp_sec_ayudas_y_subvenciones/emp_sec_ays_investigacion_desarrollo_y_tecnologia

Técnicas para el análisis del sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength, Tomás Baviera /www.riunet.upv.es

Alberto Calvo Madurga titulado Análisis de sentimientos y emociones en redes sociales usando ML presentado en la Universidad de Valladolid

Tang, D., Qin, B., Liu, T., & Yang, Y. (2015). Document modeling with gated recurrent neural network for sentiment classification

Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning word vectors for sentiment analysis

Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners

Quintero, Y. C., & García, L. A. ESTUDIO DEL ANÁLISIS DE SENTIMIENTOS BASADO EN ASPECTOS.

Cabada, R. Z., Estrada, M. L. B., & López, H. M. C. (2019). Reconocimiento multimodal de emociones orientadas al aprendizaje. *Research in Computing Science*, 148(7), 153-165.

Rodríguez Tapia, S., & Camacho-Cañamón, J. (2018). Los métodos de aprendizaje automático supervisado en la clasificación textual según el grado de especialización.

Morales Rodríguez, J. (2021). Evaluación de métodos auto-supervisados y semi-supervisados para la extracción de características visuales en el contexto de recuperación de imágenes basada en dibujos.

Análisis de Sentimientos a través de Twitter 2019 presentado en la Universidad Internacional de La Rioja, David Montoya Ruiz

Asignatura: Estrategia y Diseño de la Organización, Tema 1, J Capó Vicedo, 2021

Asignatura: Estrategia y Diseño de la Organización, Tema 2, J Capó Vicedo, 2021

Asignatura: Estrategia y Diseño de la Organización, Tema 3, J Capó Vicedo, 2021

Asignatura: Estrategia y Diseño de la Organización, Tema 4, J Capó Vicedo, 2021

Asignatura: Estrategia y Diseño de la Organización, Tema 5, J Capó Vicedo, 2021

Asignatura: Estrategia y Diseño de la Organización, Tema 6, J Capó Vicedo, 2021

Asignatura: Estrategia y Diseño de la Organización, Tema 7, J Capó Vicedo, 2021

Trabajo Fin de Grado ADE Álvaro Soler García 2021

<https://twitter.com/es/privacy> Política de Privacidad de Twitter

<https://github.com/frantrucco/BERT-Sentiment-Analysis-Twitter-Spanish>

<https://github.com/alvaroruizp/spanish-twitter-sentiment-analyzer>

<https://github.com/cardiffnlp/tweeteval>

<https://arxiv.org/abs/2104.12250>

<https://github.com/cardiffnlp/xlm-t>

<https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest?text=Covid+cases+are+increasing+fast%21>

<https://medium.com/@theclickreader/sentiment-analysis-using-python-and-deep-learning-in-3-lines-of-code-722db66fff37>

<https://neptune.ai/blog/sentiment-analysis-python-textblob-vs-vader-vs-flair>

<https://towardsdatascience.com/sentiment-analysis-in-10-minutes-with-rule-based-vader-and-nltk-72067970fb71>

<https://medium.com/analytics-vidhya/sentiment-analysis-of-social-media-data-7eb4924efe46>

<https://www.analyticsvidhya.com/blog/2021/10/sentiment-analysis-with-textblob-and-vader/>

<https://huggingface.co/blog/sentiment-analysis-twitter>

https://github.com/l33pif/Tutoriales/blob/main/Sentiment_analisis/Senitment_analysis.ipynb

Google imágenes

dreamstime.com, interfaz Instagram

14.ANEXOS

ANEXO 1: Índice de contenidos ilustrativos

Índice de gráficos

Gráfico 1: Gráfico 1: Evolución del precio medio final anual de la electricidad en España

Gráfico 2: Evolución del desempleo por trimestres en España

Gráfico 3: Porcentaje de hogares equipados con teléfono móvil en España.

Gráfico 4 : Gráfico de resultados

Gráfico 5 : Muestra de análisis individual.

Índice de cuadros

Cuadro 1: Amenazas y oportunidades

Cuadro 2: Resumen análisis funcional

Cuadro 3: Análisis DAFO

Cuadro 4: Product Backlog

Cuadro 5: Sprint 1

Cuadro 6: Sprint 2

Cuadro 7: Sprint 3

Cuadro 8: Sprint 4

Índice de esquemas

Esquema 1: Matriz de Porter

Índice de imágenes:

Imagen 1 : Metodología Scrum

Imagen 2 : Metodología en cascada

Imagen 3 :Interfaz de ejemplo de Facebook

Imagen 4 :Interfaz de ejemplo de Instagram

Imagen 5 :Interfaz de ejemplo de Twitter

Imagen 16 :Interfaz de bienvenida

Imagen 17 :Interfaz de búsqueda

Imagen 18 :Interfaz de selección

Imagen 19 :Interfaz de resultados

Índice de fragmentos de código:

Código 1 :Importación de librerías

Código 2 :Conexión con la API.

Código 3 : Solicitud de parámetros

Código 4 :Extracción por nick

Código 5 : Clasificador

Código 6 : Análisis de sentimientos

Código 7 : Generación de Gráficos

Índice de diagramas:

Diagrama 1 :Importación de librerías.

Diagrama 2 :Autenticación.

Diagrama 3 :Búsqueda.

Diagrama 4 :Limpieza del mensaje.

Diagrama 5 :Asignación y agrupación.

Diagrama 6 :Muestreo de resultados.