
**Desarrollo de una plataforma de medición instantánea
para la prevención de conductas autolesivas en
adolescentes**

**Development of an instant measurement platform for the
prevention of self-injurious behavior in adolescents**



**Trabajo de Fin de Máster
Curso 2023–2024**

Autor
Aldair Fredy Maldonado Honores

Director
Gonzalo Méndez
Pablo Gervás Gómez-Navarro

Colaborador
Colaborador 1
Colaborador 2

Máster en Ingeniería Informática
Facultad de Informática
Universidad Complutense de Madrid

Desarrollo de una plataforma de medición instantánea para la prevención de conductas autolesivas en adolescentes

Development of an instant measurement platform for the prevention of self-injurious behavior in adolescents

Trabajo de Fin de Máster en Ingeniería Informática
Departamento de Ingeniería del Software e Inteligencia Artificial

Autor
Aldair Fredy Maldonado Honores

Director
Gonzalo Méndez
Pablo Gervás Gómez-Navarro

Colaborador
Colaborador 1
Colaborador 2

Convocatoria: Junio 2024
Calificación: *Nota*

Máster en Ingeniería Informática
Facultad de Informática
Universidad Complutense de Madrid

DIA de MES de AÑO

Dedicatoria

A mi familia, quienes me han brindado su apoyo en todo momento y me han animado a seguir con el proyecto y el máster durante estos años.

Agradecimientos

Quisiera agradecer a Enrique Martín Martín, profesor y director del máster de Ingeniería Informática, por ayudarme en los momentos dónde me había atascado y que se salían del tema principal del TFM.

También quisiera agradecer a mis directores del TFM, Gonzalo Méndez y Pablo Gervás, por intentar ayudarme con mis problemas del TFM, a pesar de las dificultades externas que se han presentado en el proyecto y que nos han afectado negativamente o todos.

Resumen

Desarrollo de una plataforma de medición instantánea para la prevención de conductas autolesivas en adolescentes

El problema del suicidio y las conductas autolesivas entre los jóvenes es una preocupación de salud que ha ganado creciente atención en los últimos años, gracias a que progresivamente ha habido mayor conciencia social sobre este tema.

Es por eso que su prevención se ha vuelto un verdadero reto entre los profesionales de salud.

Bajo esta premisa surge el proyecto SIVARIA, que busca detectar de cuanto antes posibles conductas autolesivas suicidas y no suicidas entre jóvenes adolescentes de entre 12 y 21 años.

Para ello, se propone desarrollar una aplicación móvil y web que realizará una serie de cuestionarios a los usuarios sobre salud mental. Con estos datos se busca predecir posibles conductas autolesivas y que los profesionales tomen las medidas pertinentes para su tratamiento. Para realizar esta función, la aplicación se conectará internamente a un Sistema Experto que recibirá conjuntos de datos como datos de entrada, y devolverá una predicción de una posible conducta autolesiva que pueda tener el usuario. Estos datos estarían a disposición de profesionales (psicólogos, médicos, etc.), y, de forma más general, para los familiares.

Palabras clave

Sistema Experto, Python, React, React Native, red de Bayes, conjunto de datos, conductas autolesivas suicidas, conductas autolesivas no suicidas

Abstract

Development of an instant measurement platform for the prevention of self-injurious behavior in adolescents

The issue of suicide and self-harming behaviors among young people is a growing health concern in recent years, thanks to an increasing social awareness of this issue. This is why its prevention has become a significant challenge for healthcare professionals.

Under this premise, the SIVARIA project emerges, aiming to detect as early as possible both suicidal and non-suicidal self-harming behaviors among adolescents aged 12 to 21.

To achieve this goal, the project proposes the development of a mobile and web application that will administer a series of mental health questionnaires to users. Using this data, the objective is to predict potential self-harming behaviors, enabling professionals to take relevant measures for intervention. To perform this function, the application will internally connect to an Expert System that will receive datasets as input and return a prediction of a possible self-harming behavior the user may exhibit. These data would be accessible to professionals (psychologists, doctors, etc.) and, more broadly, to family members.

Keywords

Expert System, Python, React, React Native, Bayes network, dataset, suicide self-harm behaviors, non-suicide self-harm behaviors

Índice

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	3
2. Introduction	5
2.1. Motivation	5
2.2. Goals	7
3. Estado del Arte	9
3.1. Comparación de apps similares	9
3.1.1. Prevensuic	9
3.1.2. Suicide Safety Plan	9
3.1.3. ReMinder	9
3.1.4. ISNISS	10
3.1.5. PAPAGENO	10
4. Herramientas de desarrollo y tecnologías	13
4.1. Herramientas	13
4.1.1. Visual Studio Code	13
4.1.2. Github	14
4.1.3. Latex	14
4.1.4. Draw.io	14
4.2. Tecnologías	15
4.2.1. Python	15
4.2.2. React	16
5. Descripción del Trabajo	17
5.1. Investigación	17
5.1.1. Sistema Experto	17
5.1.2. Redes de Bayes	18
5.1.3. Distribuciones de probabilidad condicional	20
5.2. Desarrollo	21
5.2.1. Datasets	21
5.2.2. Desarrollo del modelo	21

5.2.3. Entrenamiento del modelo	21
5.2.4. Medición del rendimiento del modelo	21
5.2.5. Anonimización de los datos	22
6. Conclusiones y Trabajo Futuro	23
6.1. Conclusiones	23
6.2. Trabajo a futuro	23
7. Conclusions and Future Work	25
7.1. Conclusions	25
7.2. Future work	25
Bibliografía	27
A. Título del Apéndice A	29
A.1. Instrucciones para ejecutar el script	29

Índice de figuras

1.1. Causas de muertes externas en España 2008-2022. Fuente: INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	1
1.2. Causas de muertes externas en España 2013-2022 entre los jóvenes 15-19 años. Fuente: INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	2
1.3. Causas de muertes externas en España 2013-2022 entre los jóvenes 20-24 años. Fuente: INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	3
1.4. Número de suicidios en España entre los años 2000 y 2022. Fuente: Fundación española para la prevención del suicidio, Estadísticas de Defunción por Causa de Muerte 2022. INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	4
2.1. Causes of external death in Spain from 2008 to 2022. Source: INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	5
2.2. Causes of external death in Spain from 2013 to 2022 among young people aged 15 and 19 years old. Source: INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	6
2.3. Causes of external death in Spain from 2013 to 2022 among young people aged 20 and 24 years old. Source: INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	7
2.4. Number of suicides in Spain between 2000 and 2022. Source: Fundación española para la prevención del suicidio, Statistics of Causes of Death in 2022 of people of all ages. INE, 2024: https://www.ine.es/jaxiT3/Tabla.htm?t=7947	8
3.1. Interfaces de Prevensuic	10
3.2. Interfaces de Suicide Safety Plan.	10
4.1. Interfaz web de Draw.io	15
5.1. Estructura del Sistema Experto	18
5.2. Página inicial del sitio web de SIVARIA	18
5.3. Encuesta a jóvenes de más de 16 años de SIVARIA	19
5.4. Ejemplo de un Diagrama de Bayes. Fuente, Javatpoint	19

Índice de tablas

3.1.	Tabla comparativa de aplicaciones de prevención contra el suicidio. Última actualización: mayo de 2024	11
3.2.	Continuación de la Tabla 3.1	11
4.1.	Popularidad de los lenguajes de programación, Fuente: PYPL (2023)	13
5.1.	CPD de la variable <i>Alarm</i>	20

Capítulo 1

Introducción

1.1. Motivación

El suicidio es un problema de salud a nivel global (Organization et al. (2014)), siendo la principal causa de muerte no natural entre jóvenes y adultos (Al-Halabí y Fonseca-Pedrero (2021), Gore et al. (2011)). Este tema del suicidio ha ido adquiriendo cada vez más atención dentro del ámbito de la salud, y su prevención ha presentado un reto para todos los profesionales, ya que dicha problemática afecta tanto a los propios individuos, como también a los familiares y amigos cercanos.

Desgraciadamente, España no ha sido la excepción. Según las estadísticas del INE, el suicidio sigue siendo la principal causa de muerte externa en España desde 2008 (1.1) y una de las principales de los jóvenes de entre 15 y 24 años (Figuras 1.2 y 1.3)(Eterovic).

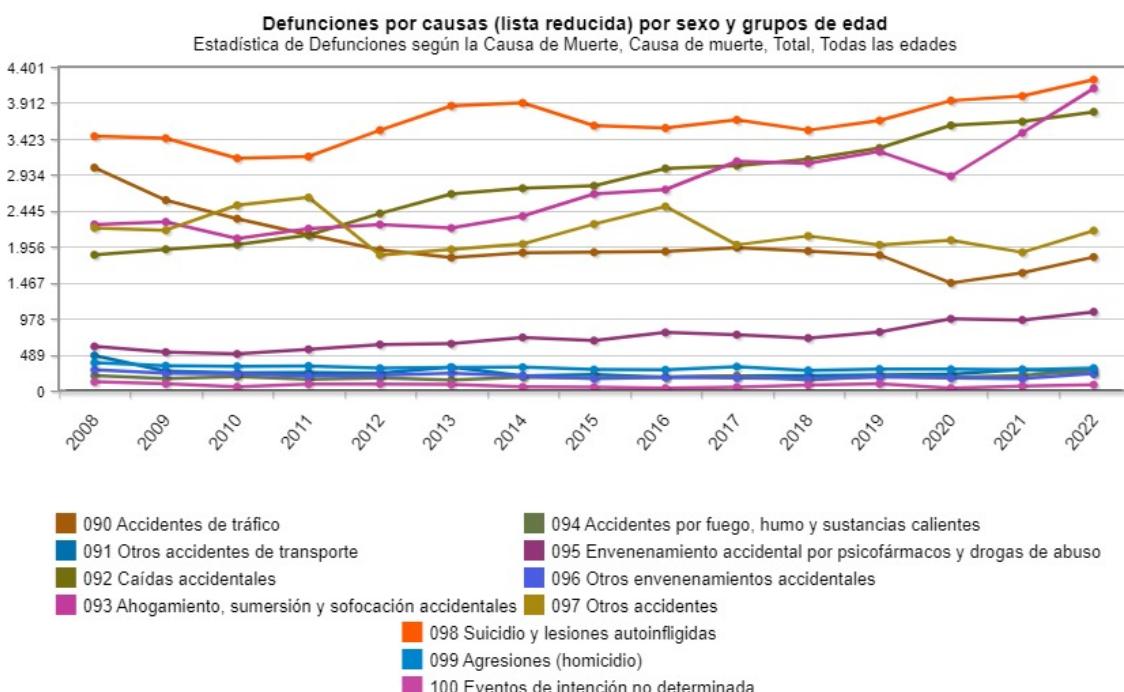


Figura 1.1: Causas de muertes externas en España 2008-2022. Fuente: INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

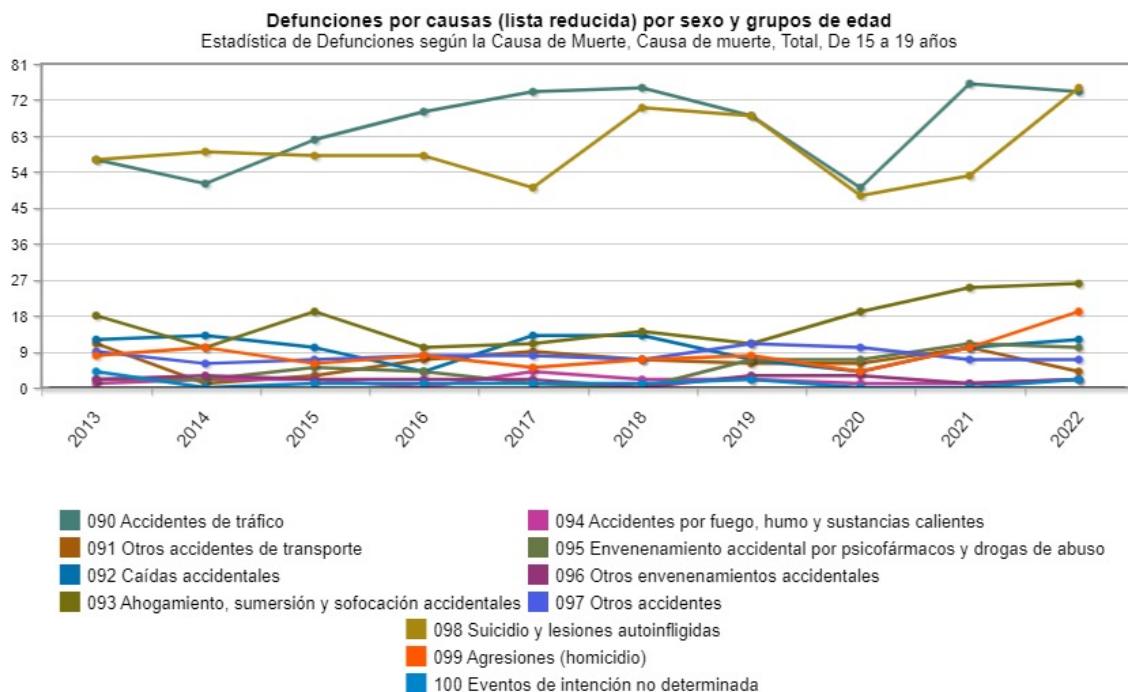


Figura 1.2: Causas de muertes externas en España 2013-2022 entre los jóvenes 15-19 años.
Fuente: INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

No sólo eso, sino que además el número de suicidios en España ha aumentado desde 2018, marcando cifras históricas (1.4).

A estas estadísticas también se le acompañan encuestas realizadas a grupos de jóvenes españoles, como el que se llevó a cabo por Muela et al. (2024).

Esto evidencia un claro aumento de suicidios y de las conductas autolesivas en los adolescentes en España.

Es por ello que en los últimos años, se están centrando los esfuerzos y los recursos económicos en la creación de programas de prevención del suicidio. En estos planes, es fundamental realizar diagnósticos lo más pronto posible para detectar a tiempo síntomas mentales en la persona, y por lo tanto, poder ofrecerle un tratamiento. El problema es que muchas veces, las personas con riesgo de adoptar estas conductas suicidas y autolesivas no son conscientes de que los padecen o que pueden padecerlos, por lo que no acuden a los planes de prevención. Esto trae, como consecuencia, el agravamiento del problema, hasta acabar en un desenlace fatal. Es importante el poder informar a la persona a tiempo sobre su estado de salud mental de forma automática, rápida y eficaz, para que así sea consciente del riesgo al que está sometido, y por lo tanto, pueda pedir ayuda con anticipación.

Ante este panorama, se busca desarrollar herramientas para prevenir este tipo de conductas en la juventud española. Es aquí donde comienza a jugar un papel fundamental la inteligencia artificial.

De acuerdo a al artículo de Fonseka et al. (2019), la inteligencia artificial se está utilizando cada vez más en el campo de la salud mental, siendo de gran utilidad para la prevención del suicidio y dar soporte a los profesionales para conseguir una evaluación exacta del riesgo de suicidio. De esta manera, se pueden tratar a los potenciales pacientes que posean estas conductas autolesivas o no autolesivas en etapas muy tempranas del desarrollo de estos comportamientos.

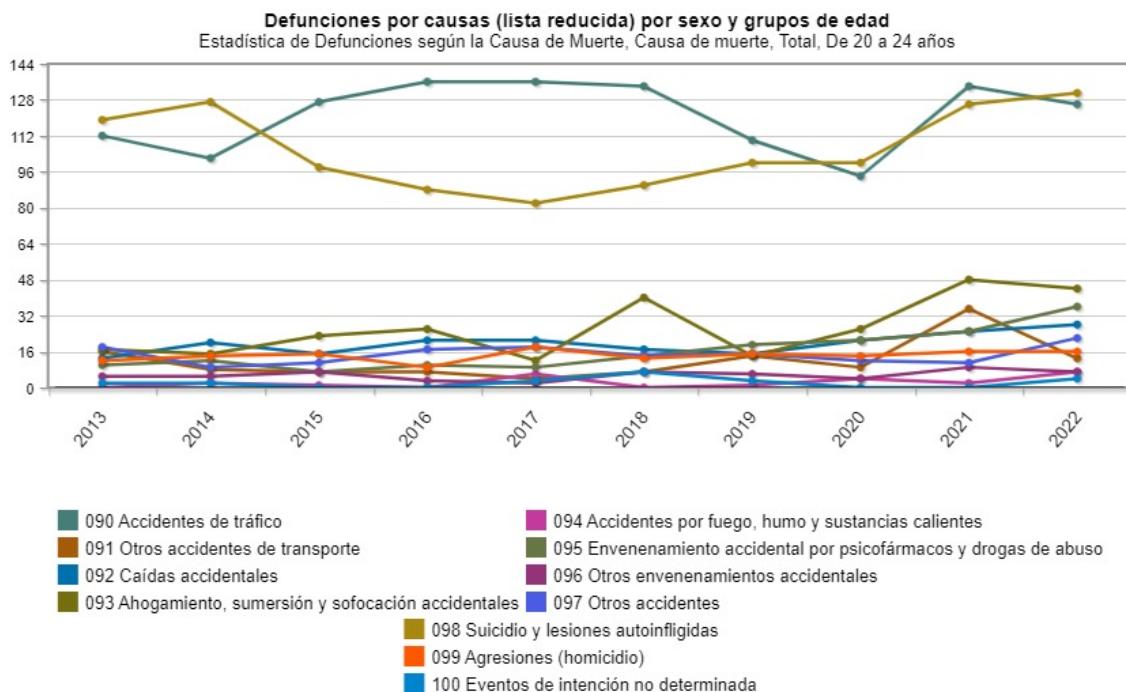


Figura 1.3: Causas de muertes externas en España 2013-2022 entre los jóvenes 20-24 años.
Fuente: INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

1.2. Objetivos

Es precisamente por este aumento del uso de la IA en el campo de la salud en estos últimos años, se ha propuesto este Trabajo de Fin de Máster. El objetivo es desarrollar un algoritmo para medir y prevenir las conductas autolesivas suicidas y no suicidas en jóvenes adolescentes.

Para ello, se deben realizar una serie de formularios hechos por profesionales a los jóvenes. Estas respuestas serán tratadas de forma anónima y se enviarán al modelo que se encargará de predecir posibles conductas autolesivas y no autolesivas que podría tener el usuario en base a las respuestas obtenidas del formulario.

Para ello, se propone la implementación de un Sistema Experto que se encargue del proceso de predicción, diagnóstico y evaluación, simulando así las decisiones que tomaría un profesional real.

Con estas premisas, se intentan fijar esta serie de objetivos:

- Diseñar un modelo inicial de Sistema Experto.
- Pasar el diseño del modelo a código.
- Entrenar al modelo con una serie de datasets proporcionados por parte de los profesionales involucrados en este proyecto.
- Testear el modelo con otros datos de prueba. Se cuantifica el rendimiento del modelo, se mide y en base a los resultados obtenidos, se calibra el modelo. Más adelante se explicará las distintas formas de medir el rendimiento de un modelo probabilístico.
- Se repite el proceso de entrenamiento y testeo del modelo hasta obtener un resultado

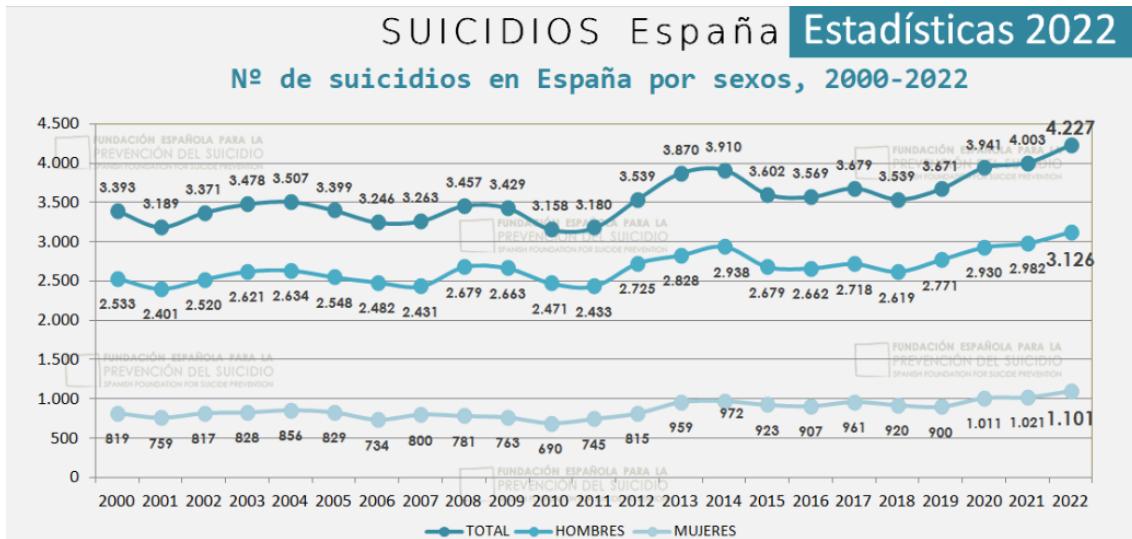


Figura 1.4: Número de suicidios en España entre los años 2000 y 2022. Fuente: Fundación española para la prevención del suicidio, Estadísticas de Defunción por Causa de Muerte 2022. INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

lo suficientemente óptimo para que el modelo pueda realizar predicciones de casos reales.

Capítulo 2

Introduction

2.1. Motivation

Suicide is a global health problem (Organization et al. (2014)), being the leading cause of unnatural death among young people and adults (Al-Halabí y Fonseca-Pedrero (2021), Gore et al. (2011)). The issue of suicide has been gaining increasing attention within the health field, and its prevention has been a challenge for health professionals, as the problem affects both the individuals themselves, as well as their family members and close friends.

Unfortunately, Spain has been no exception. According to INE statistics, suicide continues to be the leading cause of external death in Spain since 2008 (2.1) and one of the main ones among young people aged 15 and 24 years old (Figures 2.2 and 2.3)(Eterovic).

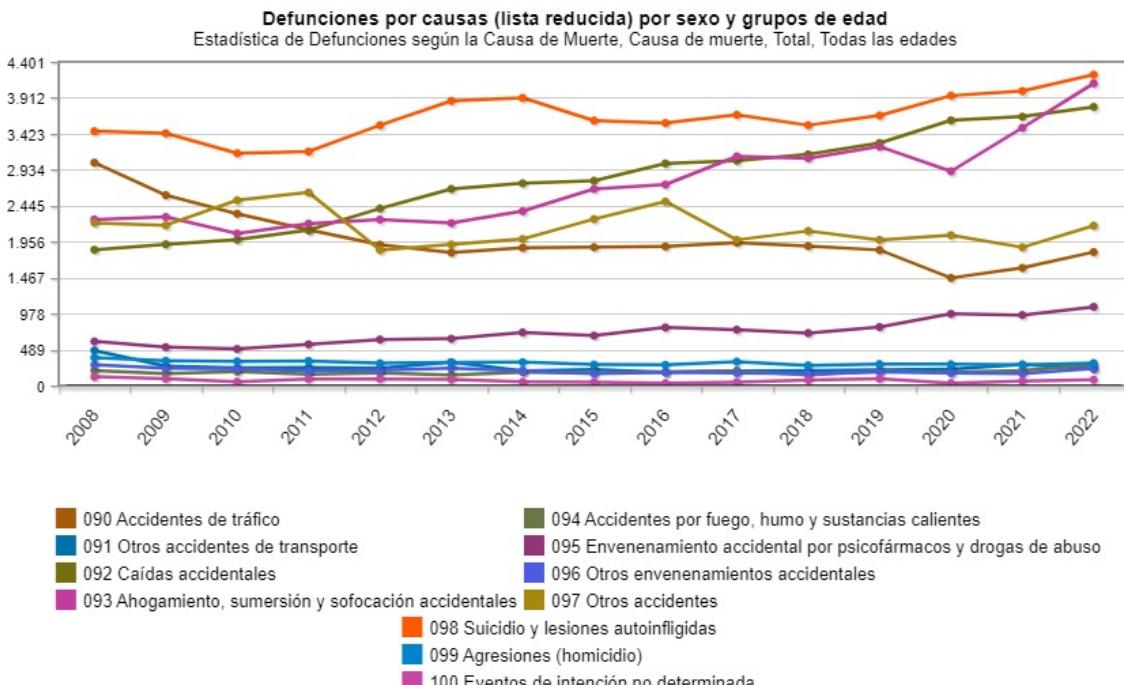


Figura 2.1: Causes of external death in Spain from 2008 to 2022. Source: INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

Furthermore, the number of suicides in Spain has increased since 2018, marking histo-

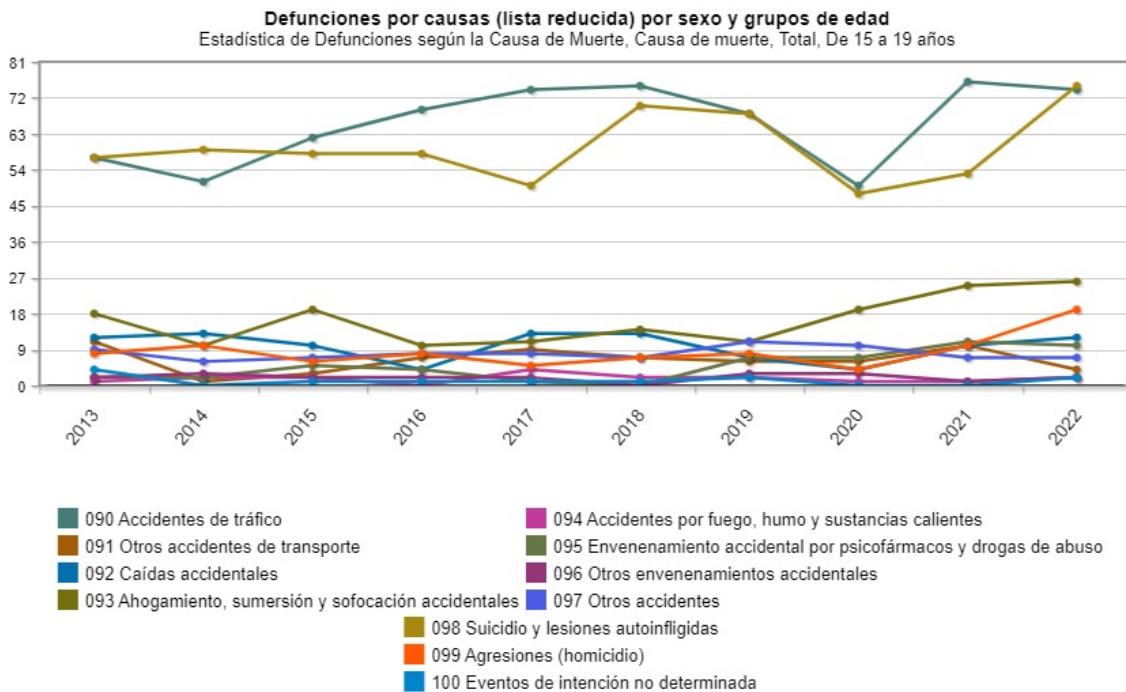


Figura 2.2: Causes of external death in Spain from 2013 to 2022 among young people aged 15 and 19 years old. Source: INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

rical figures (2.4).

These statistics are also accompanied by surveys conducted among groups of young Spaniards, such as the one conducted by Muela et al. (2024).

This provides evidence of a clear increase in suicides and self-injurious behaviors among adolescents in Spain.

For this reason, in recent years, efforts and economic resources have been focused on the creation of suicide prevention programs.

In these plans, it is essential to make diagnoses as early as possible to detect mental symptoms in the person, and therefore, to be able to offer treatment. The problem is that people at risk of adopting these suicidal and self-injurious behaviors are often unaware that they suffer from them or that they may suffer from them, so they do not attend prevention plans. As a consequence, this leads to a worsening of the problem, until it ends in a fatal outcome. It is important to be able to inform the person in time about his or her mental health condition automatically, quickly and effectively, so that he or she is aware of the risk to which he or she is subjected, and therefore can ask for help in advance.

Against this background, it is essential to develop strategies, tools and methods to prevent this type of behaviour among Spanish youth. This is where artificial intelligence begins to play a fundamental role.

According to the following article (Fonseka et al. (2019)), artificial intelligence is being used progressively in the field of mental health, being useful in suicide prevention and providing support to professionals to achieve an effective and efficient diagnosis and assessment. In this way, potential patients with these self-injurious and non-self-injurious behaviours can be treated in early stages.

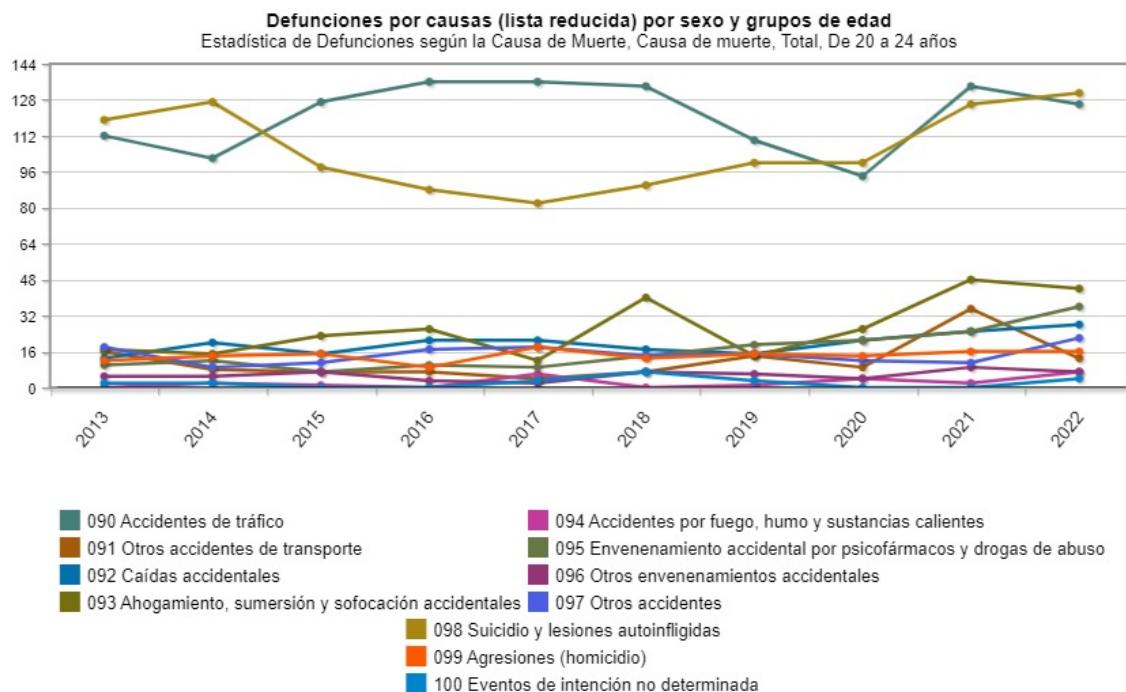


Figura 2.3: Causes of external death in Spain from 2013 to 2022 among young people aged 20 and 24 years old. Source: INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

2.2. Goals

It is precisely because of this increase in the use of AI that it has been proposed to develop an algorithm to measure and prevent suicidal and non-suicidal self-injurious behaviour in young adolescents.

To do this, a series of forms are to be filled out by professionals with young people. These answers will be treated anonymously and will be sent to the model that predicts the possible self-harming and non-self-harming behaviours that the user could have based on the answers obtained from the form.

In order to create this model, the implementation of an Expert System was proposed. This Expert System will be in charge of the prediction, diagnosis and evaluation process, thus simulating the decisions that a real professional would take.

Based on these premises, the following initial individual goals were set:

- To design an initial model of the Expert System. The types of systems that exist and the one selected for this project will be explained later on.
- To translate the design to code. A programming language and a set of libraries will be selected to implement the model.
- To train the model with datasets provided by the experts that are involved in this project.
- To test the model with part of the provided. The testing will be quantified in a score that will be based on the results obtained from the testing. This score will be interpreted in order to apply possible improvements to the model. The different methods to obtain this score will be explained later on.

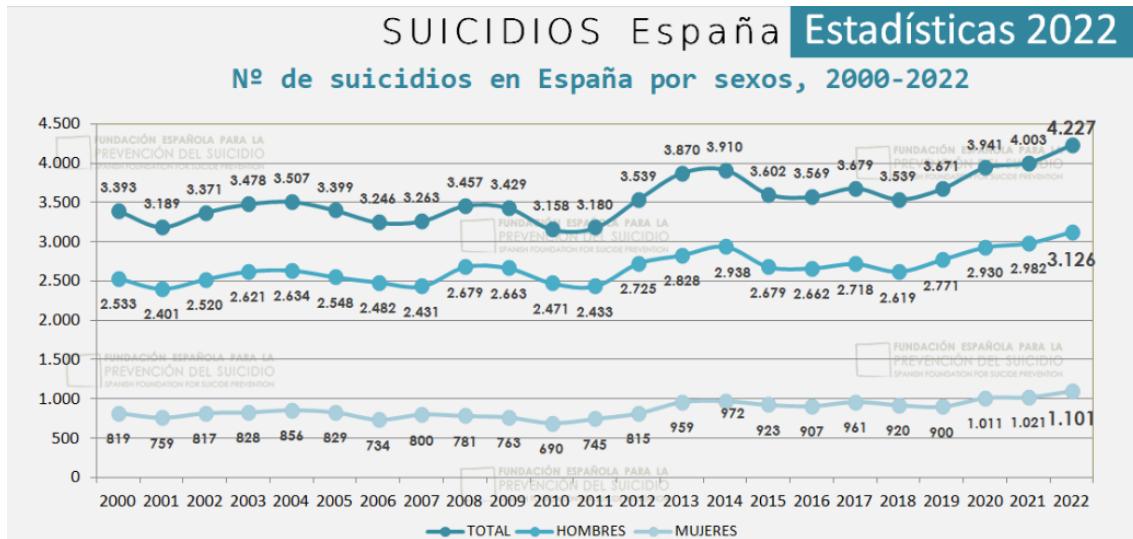


Figura 2.4: Number of suicides in Spain between 2000 and 2022. Source: Fundación española para la prevención del suicidio, Statistics of Causes of Death in 2022 of people of all ages. INE, 2024: <https://www.ine.es/jaxiT3/Tabla.htm?t=7947>

- To Repeat training and testing process until getting an acceptable score from the model. That means that it is good enough to make predictions in real cases.

Capítulo 3

Estado del Arte

En este capítulo se expondrá la importancia del uso de la inteligencia artificial en el campo de la salud mental, así como ejemplos y referencias relacionados con el proyecto.

3.1. Comparación de apps similares

Antes de desarrollar la aplicación, se entregaron un documento con una lista de aplicaciones analizadas por los directores que se están desarrollando en este campo.

De la lista completa de aplicaciones, se han desatulado cinco de ellas, ya que eran las que estaban más relacionadas con las intenciones de nuestra aplicación, además de que eran las que proporcionaban una descripción más completa de sus funcionalidades y objetivos.

3.1.1. Prevensuic

Prevensuic¹ es un programa creado por la *Fundación Española para la Prevención del Suicidio*, destinado a la prevención, divulgación y la formación acerca del suicidio. Dicho programa cuenta con un sitio web y una aplicación móvil (disponible tanto en Android como en iOS) para ser accesible a un mayor número de personas, tanto adultos como jóvenes.

3.1.2. Suicide Safety Plan

Suicide Safety Plan² es una aplicación móvil desarrollada por Inquiry Health LLC y está disponible tanto en Android como en iOS. La aplicación es una especie de aplicación recordatorio para que el usuario pueda establecer qué cosas pueden afectar a su salud mental, e incluye información para ayudar a la persona a evaluar su propio estado mental. Además, incluye una serie de teléfonos de contacto en caso de necesitar ayuda.

3.1.3. ReMinder

ReMinder³ es un programa inglés dirigido a gente adulta y que consta de un cuestionario de 10 preguntas.

¹<https://www.prevensuic.org/>

²https://play.google.com/store/apps/details?id=com.moodtools.crisis.app&hl=en_US

³<https://www.emhprac.org.au/directory/reminder-suicide-safety-plan/>



Figura 3.1: Interfaces de Prevensuic

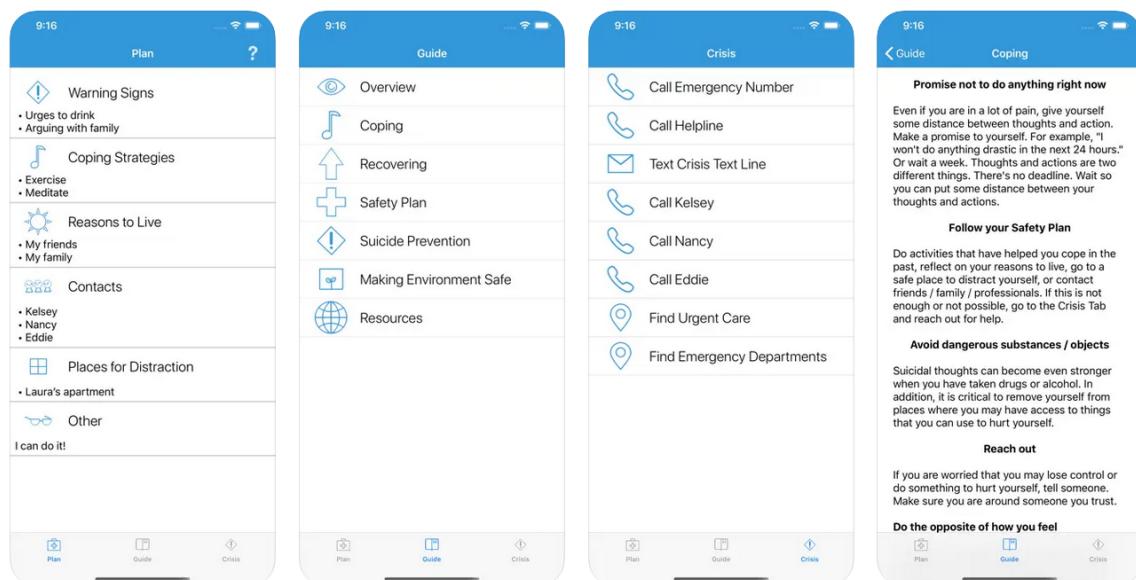


Figura 3.2: Interfaces de Suicide Safety Plan.

3.1.4. ISNISS

En este caso, ISNISS⁴ es un sitio web que se encarga de dar información sobre cómo realizar una evaluación correcta del riesgo de suicidio en un paciente, aunque no la realiza el propio sitio web, aparte de otra serie de recomendaciones.

3.1.5. PAPAGENO

Similar a ISNISS, PAPAGENO⁵ es un blog que ofrece información sobre la postvención y recomendaciones para la evaluación por los sanitarios.

Haciendo una comparación entre todas ellas, se llegó a las siguientes tablas 3.1 y 3.2. De todas las aplicaciones de las Tablas 3.1 y 3.2, uno de los más destacados es la

⁴<https://www.isniss.es/>

⁵<https://papageno.es/about/quienes-somos>

	Prevensuic	Suicide Safety Plan	ReMinder
Tipo de aplicación	Web y Móvil	Móvil	Móvil
Población objetivo	Adultos	Adultos	Adolescentes
Idioma	Español	Inglés	Inglés
Evaluación del riesgo	X	X	X
Estrategias de afrontamiento	✓	✓	X
Psicoeducación	✓	✓	✓
Afectados	✓	✓	✓
Profesionales	✓	X	X
Familiares	✓	X	X

Tabla 3.1: Tabla comparativa de aplicaciones de prevención contra el suicidio. Última actualización: mayo de 2024

	ISNISS	PAPAGENO
Tipo de aplicación	Web	Web
Población objetivo	Adolescentes	Adultos
Idioma	Español	Español
Evaluación del riesgo	X	X
Estrategias de afrontamiento	-	-
Psicoeducación	X	-
Afectados	X	✓
Profesionales	X	✓
Familiares	X	✓

Notas: " - ": no se sabe si la característica está en la aplicación.

Tabla 3.2: Continuación de la Tabla 3.1

aplicación de Prevensuic.

Una de las aplicaciones que más fueron analizadas fue Prevensuic. Se le considera una aplicación o programa muy completa y psicoeducativo, es decir, que es capaz de brindar información a las personas que están sufriendo de un trastorno psicológico.

Sin embargo, Prevensuic presenta dos grandes inconvenientes:

- La aplicación no es capaz de realizar una evaluación del riesgo de suicidio.
- La versión de móvil de Prevensuic no está disponible para las últimas versiones de (Android 14, por ejemplo). Esto se debe a que no se actualiza de manera recurrente. Su última actualización data del 11 de abril de 2019. Esto supone una disminución considerable de la accesibilidad de la aplicación y, por lo tanto, del programa en general.

Desde SIVARIA, se busca proporcionar una herramienta web y móvil que pueda proporcionar una evaluación automática del riesgo de suicidio en base a las respuestas de un cuestionario. Dicha aplicación se espera que reciba actualizaciones frecuentemente, además de que estará dirigida principalmente a jóvenes, pero también pueden acceder profesionales y los familiares y conocidos del paciente.

Capítulo 4

Herramientas de desarrollo y tecnologías

4.1. Herramientas

4.1.1. Visual Studio Code

Visual Studio Code es un editor de código fuente, gratuito, de código abierto y multiplataforma, ya que fue creado por Microsoft para Windows, Linux, MacOS y Web. Cuenta con herramientas internas de depuración, resaltado de sintaxis, control de versiones de Git, autocompletado de código, entre otros.

Se ha propuesto que se utilizará esta herramienta para desarrollar el prototipo de la aplicación en React.

Cabe resaltar que, en un principio, se planeaba utilizar Android Studio como entorno de desarrollo, y programar la aplicación con Java, por su popularidad en el desarrollo de aplicaciones web y móviles, como se puede observar en la Tabla 4.1.

Sin embargo, Android Studio facilitaba exclusivamente el desarrollo de aplicaciones compatibles con dispositivos móviles Android en Java. Por ende, si se hubiera persistido con esta propuesta, habría existido el riesgo de que la aplicación presentara errores durante la ejecución o, incluso, que no pudiera ser ejecutada en otros sistemas operativos distintos, como iOS.

Otra de las razones por las que finalmente se ha optado por el uso de VS Code fue para poder desarrollar scripts en Python. Al ser de código abierto, ha permitido que se puedan

Rank	Lenguaje de programación	Cuota de mercado
1	Python	28.43 %
2	Java	16.04 %
3	JavaScript	8.72 %
4	C/C++	6.65 %
5	C#	6.63 %
6	R	4.63 %
7	PHP	4.45 %
8	TypeScript	2.96 %
9	Swift	2.71 %
10	Rust	2.53 %

Tabla 4.1: Popularidad de los lenguajes de programación, Fuente: PYPL (2023)

desarrollar diferentes plugins open-source para facilitar el desarrollo de Python y React en este editor.

4.1.2. Github

Todas las novedades y cambios en el código del Sistema Experto, aplicación o en la documentación se añaden a un repositorio de Github. Github es una plataforma en línea para alojar una gran diversidad de proyectos. El enlace al repositorio de SIVARIA es el siguiente:

<https://github.com/NILGroup/TFM-2324-SIVARIA>

Entre las características más notables de Github, destaca su sistema de control de versiones basado en Git. Git facilita el seguimiento de todos los cambios realizados en el código fuente, abarcando todas las ramas del repositorio. Esta capacidad resulta sumamente útil para administrar el desarrollo del software de manera eficiente y efectiva (Wikipedia (2024a)).

Con respecto al flujo de trabajo, se tenía pensado seguir el flujo de trabajo *GitFlow* que ya está implementado en Github, en donde se parte de una *main*, que simulará ser una rama de producción. A partir de esta rama, deriva otra nueva rama *develop*. Finalmente, a partir de esta rama derivan una serie de subramas llamadas *feature*, una por cada tarea o corrección que surja.

De esta manera, cada tarea que se vaya a desarrollar se hará en una rama nueva *feature*, y cuando dicha tarea se termine, se mergea con la rama *develop*, y posteriormente con la rama *main*, liberando las tareas realizadas para que los directores o cualquier usuario pueda testear las nuevas implementaciones.

Siguiendo esta metodología, podemos tener el trabajo bien diferenciado por cada tarea a realizar, además de que evita el riesgo de presentar a los usuarios finales código defectuoso o con errores.

Por otro lado, en local se utilizará Github Desktop para subir todos los cambios en local al repositorio remoto de Github.

Github Desktop es una herramienta local que proporciona una interfaz de usuario que permite a usuarios convencionales poder gestionar los repositorios remotos sin necesidad de conocer en profundidad los comandos Git necesarios.

4.1.3. Latex

Latex ((Urban, 1986)) es un sistema open-source de composición de texto ampliamente utilizado en la redacción de documentos, textos y artículos académicos y científicos. Una de las características más destacadas de Latex es que se basan en comandos de texto, a diferencia de un procesador de textos convencional, que utilizan un enfoque visual. Esto nos permite escribir caracteres especiales y funciones matemáticas y, además de enumerar más fácilmente imágenes y tablas.

Esta herramienta es la que se ha utilizado para redactar esta memoria en su totalidad.

4.1.4. Draw.io

Draw.io es una herramienta online de dibujos gráfico que es utilizado para crear diagramas de flujo, organigramas, diagramas UML o de red de forma rápida y sencilla. Es una herramienta gratuita que se puede vincular con Google Drive, de tal forma que podemos guardar el archivo draw.io en la nube. Se ha utilizado para diseñar y exportar los diagramas de Bayes en su totalidad.

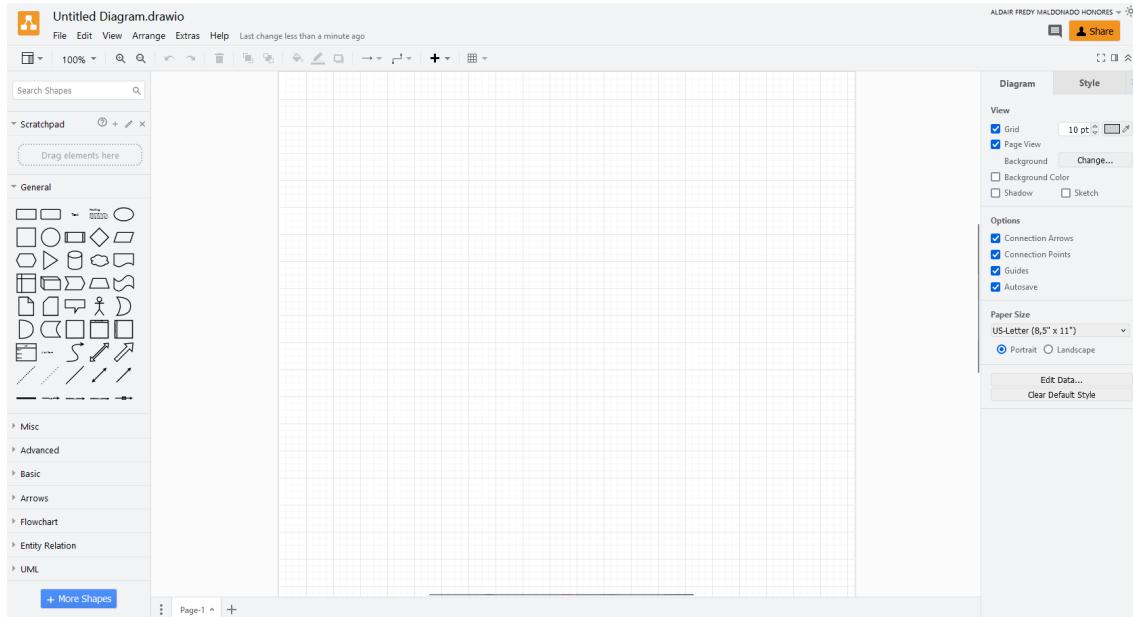


Figura 4.1: Interfaz web de Draw.io

4.2. Tecnologías

4.2.1. Python

Se decidió utilizar el lenguaje de programación Python para desarrollar el Sistema Experto de la aplicación, concretamente, la versión de Python utilizada es la 3.9 (Python (2020)).

Los motivos de su elección se deben a la gran variedad de bibliotecas de código abierto que ofrece Python para el aprendizaje automático, además de la manipulación y visualización de datos.

Por otro lado, Python es fácilmente escalable, es decir, se puede utilizar tanto en proyectos pequeños como grandes. Esto se debe a que Python da soporte a tecnologías que permiten el procesamiento distribuido, como por ejemplo, Apache Spark.

4.2.1.1. Bibliotecas utilizadas

Se han utilizado las siguientes bibliotecas de código libre de Python:

- **Pandas**: ofrece estructuras de datos para poder almacenar el contenido de los datasets. En este caso, los datos se almacenarán en dataframes, que son estructuras de datos similares a las tablas de las bases de datos SQL. Por lo tanto, nos permitirán realizar consultas, inserciones, borrados y modificaciones sobre ellas de forma bastante intuitiva.
- **Scikit-learn**: incluye algoritmos de clasificación, regresión y análisis de grupos útiles para el desarrollo y evaluación de modelos de aprendizaje automático, aunque también se utilizan en la minería y análisis de datos.
- **Matplotlib**: permite crear gráficos estáticos, interactivos y animaciones de alta calidad. Es una biblioteca muy utilizada principalmente por su capacidad para generar

una variedad de tipos de gráficos y visualizaciones, además de que permite exportarlos en formato PNG, JPEG, PDF y SVG.

Cabe destacar que antes de Scikit-learn, se estudiaron la posibilidad de utilizar otros paquetes para la gestión de la red de Bayes. Estos son pgmpy(Ankan y Panda (2015)), pybnfesian (Atienza et al. (2022)) y pomegranate(Schreiber (2014)).

En el caso de los pybnfesian y pomegranate, se descartaron por la complejidad de su sintaxis. Con respecto a pgmpy, se llegó a desarrollar una implementación alternativa del modelo en pgmpy. Sin embargo, presentaba severos problemas con la capacidad de memoria local para crear el modelo. Esto provocaba que el proceso de entrenamiento tardase demasiado tiempo, y los procesos de testeo y predicción directamente no devolvía ningún tipo de respuesta.

4.2.2. React

React es una biblioteca de Javascript de código abierto que permite crear interfaces de usuario, tanto para aplicaciones de escritorio como para teléfonos móviles.

Fue desarrollada y lanzada por Facebook en el año 2013, aunque también es usado por otras aplicaciones, como Instagram, Whatsapp, X, entre otros.

Se decidió este lenguaje por delante de otros lenguajes de desarrollo de aplicaciones, como Kotlin, por la falta de familiaridad con dicho lenguaje, o Java, por su incompatibilidad con los distintos sistemas operativos. Esto último, es una de las características más destacadas de React para nuestro proyecto. Permite que se puedan desarrollar aplicaciones móviles para Android, iOS y UWP (Wikipedia (2024b)).

Para facilitar y agilizar el desarrollo de la aplicación, se ha propuesto usar un framework de React, ya que traen varios componentes básicos implementados, como botones, enlaces, formularios, etc. para ahorrar tiempo y esfuerzo. Existen diversos framework de React, cada una enfocada en un propósito específico.

- **React Native:** ampliamente utilizado en aplicaciones móviles, ya que permite a los desarrolladores construir aplicaciones nativas para iOS y Android utilizando React.js y JavaScript.
- **React Bootstrap:** es un framework que implementa a su vez el framework CSS de Bootstrap, que trae de por sí un abanico de estilos previamente diseñados y que el desarrollador sólo necesita declararlos dentro del código.
- **Next.js:** se utiliza para construir aplicaciones web de gran tamaño. Una de sus características destacables es que ofrece funcionalidades de enrutamiento y renderización del lado del servidor (SSR) y la generación de sitios estáticos.
- **Gatsby:** es un framework que usa GraphQL para gestionar los datos de la aplicación, además de que cuenta con una gran variedad de plugins para extender sus funciones. Es especialmente popular para la creación de blogs, sitios web de comercio electrónico y portafolios.

En este caso, React Native fue seleccionado como el framework de desarrollo. El motivo principal de esta decisión es que ... - Framework ->React Native (nuevo apartado?, utilizado por empresas como Facebook, referencia: React (2015))

Capítulo 5

Descripción del Trabajo

5.1. Investigación

Teniendo en cuenta los objetivos del proyecto, el primer paso fue diseñar la infraestructura del Sistema Experto. Para ello, era necesario saber qué es un Sistema Experto, los distintos tipos que había, cuál era el tipo de sistema elegido, y finalmente, listar las herramientas y tecnologías que se requerían para su desarrollo.

5.1.1. Sistema Experto

Un sistema experto es un programa informático diseñado para simular el conocimiento y las habilidades analíticas de un especialista en un campo concreto. Su función es simular la forma en la que un especialista toma una decisión.

Dependiendo del tipo del sistema experto, pueden estar basados en:

- Reglas previamente establecidas (RBR, Rule Based Reasoning) o encadenamiento hacia atrás.
- Basados en casos (CBR, Case Based Reasoning) o encadenamiento hacia adelante.
- **Basados en redes bayesianas.**

En nuestro caso, se ha decidido implementar el Sistema Experto mediante el uso de redes bayesianas, ya que es el tipo más comúnmente utilizado para este tipo de aplicaciones con sistemas expertos o variantes basadas en la inteligencia artificial.

Con respecto al proyecto, el objetivo de este sistema es:

- Identificar y, por lo tanto, predecir, aquellos jóvenes que cumplen con las condiciones para ser considerados de alto riesgo (pronóstico).
- Monitorizar sus conductas de riesgo mediante evaluación momentánea (Ecological Momentary Assessment) (monitorización).
- Prevenir que se produzcan conductas autolesivas futuras (prevención), mediante la provisión de indicaciones y actuaciones personalizadas y ajustadas a los niveles de riesgo bajo la supervisión de profesionales.

Por otro lado, se ha propuesto que la estructura del sistema sea de tal forma que reciba una serie de datos de entrada, como se muestra en la Figura 5.1.

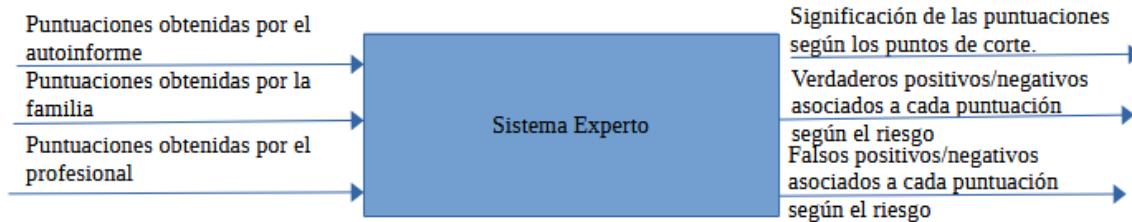


Figura 5.1: Estructura del Sistema Experto

El sistema recibirá una serie de puntuaciones obtenidas de las encuestas realizadas a los usuarios por el autoinforme, a las familias de los usuarios y a los profesionales de la salud (médicos, psicólogos, etc.).

Dichas encuestas se recogen en el sitio web de SIVARIA (Figuras 5.2 y 5.3).



Figura 5.2: Página inicial del sitio web de SIVARIA

En base a estos datos y al entrenamiento realizado, el sistema los procesará y devolverá una serie de positivos y negativos, tanto verdaderos como falsos. Esto será interpretado por la aplicación para determinar si existen conductas autolesivas en el usuario o los potenciales comportamientos que podría llegar a adoptar.

5.1.2. Redes de Bayes

Como se ha explicado anteriormente, se ha optado por implementar el Sistema Experto mediante una red de Bayes (Sucar y Tonantzintla (2006)).

Las redes bayesianas, propuesto inicialmente por Judea Pearl en 1988, son un modelo probabilístico en el que se representan las relaciones probabilísticas entre las variables (nodos) y las dependencias condicionales que hay entre ellas, formando de esta manera un grafo acíclico dirigido (DAG, *Directed Acyclic Graph*).

Para que un grafo pueda ser considerado un DAG, es necesario que todos los arcos (*edges*) del grafo sean dirigidos, es decir, que indiquen una y sólo una dirección. Además, el grafo no debe tener ciclos, es decir, que dado un vértice v , no hay un camino directo que

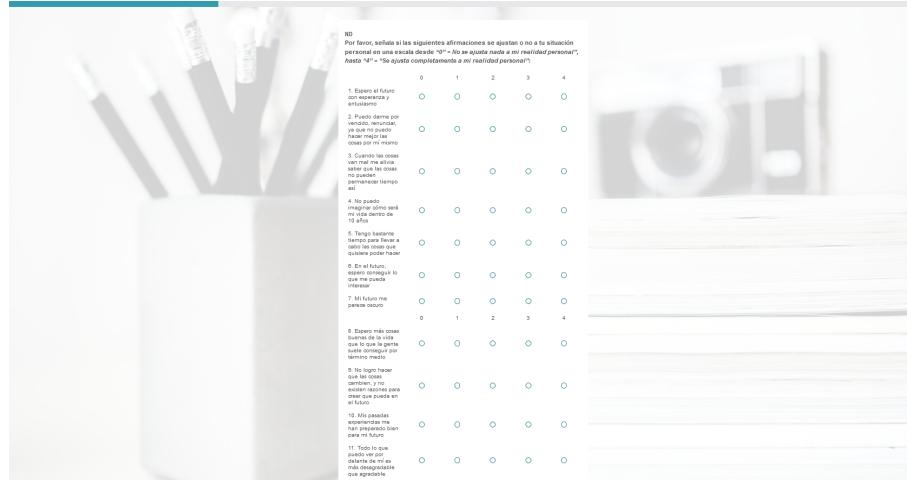


Figura 5.3: Encuesta a jóvenes de más de 16 años de SIVARIA

empiece y termine en v .

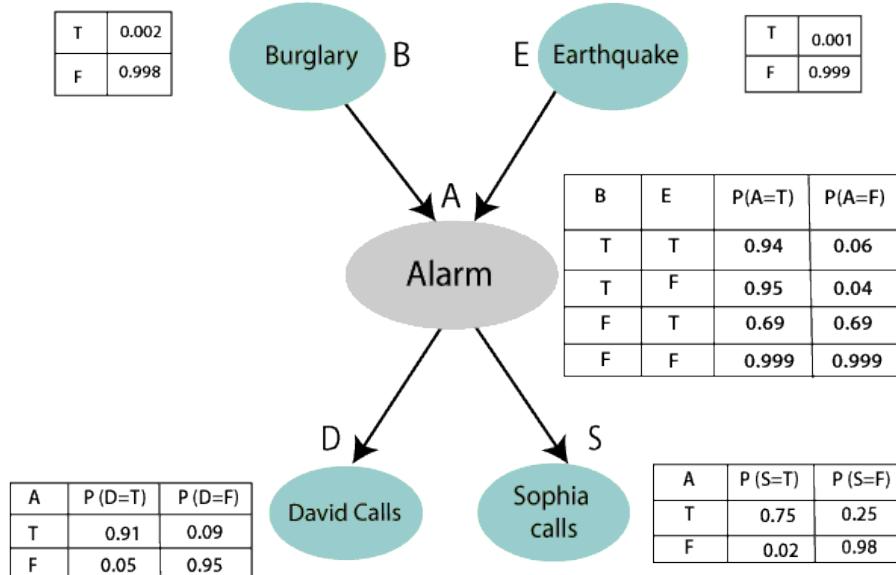


Figura 5.4: Ejemplo de un Diagrama de Bayes. Fuente, Javatpoint

Dentro de una red bayesiana, es necesario entender los conceptos de probabilidad conjunta y probabilidad condicional.

5.1.2.1. Probabilidad conjunta

La probabilidad conjunta es la probabilidad de que dos o más eventos ocurran al mismo tiempo. En una red de Bayes, la probabilidad conjunta de n variables se puede representar como:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | X_{p(i)}) \quad (5.1)$$

Donde $X_{p(i)}$ es el padre de X_i .

Por ejemplo, dado:

Burglary	Earthquake	P(Alarm=T)	P(Alarm=F)
T	T	0.94	0.06
T	F	0.95	0.04
F	T	0.69	0.69
F	F	0.999	0.999

Tabla 5.1: CPD de la variable *Alarm*

- D = *David Calls*.
- S = *Sophie Calls*.
- A = *Alarm*.
- B = *Burglary*.
- E = *Earthquake*.

La probabilidad conjunta de la Figura 5.4 sería:

$$\begin{aligned}
 P(D, S, A, B, E) &= P(D|A)P(S, A, B, E) \\
 &= P(D|A)P(S|A)P(A, B, E) \\
 &= P(D|A)P(S|A)P(A|B, E)P(B, E) \\
 &= P(D|A)P(S|A)P(A|B, E)P(B)P(E)
 \end{aligned} \tag{5.2}$$

5.1.2.2. Probabilidad condicional

La probabilidad condicional es la probabilidad de que un evento A ocurra dado otra evento B, $P(A|B)$.

Gracias a estos dos conceptos, se puede explicar uno de los elementos imprescindibles de las redes bayesianas: las tablas de distribución de probabilidad condicional.

5.1.3. Distribuciones de probabilidad condicional

Como se puede observar en la red de la Figura 5.4, una de las características de una red de Bayes son las tablas de distribución de probabilidad condicional (*CPD, Conditional Probability Distribution*).

Una CPD es una tabla donde se especifica la probabilidad de cada posible valor de una variable dada una combinación específica de los valores de sus variables padres. Tomando como ejemplo la Figura 5.4, la ecuación de la probabilidad condicional de que la alarma suene (*Alarm*) sería la siguiente:

$$P(\text{Alarm}|\text{SophiaCalls}, \text{DavidCalls}, \text{Burglary}, \text{Earthquake}) = P(\text{Alarm}|\text{Burglary}, \text{Earthquake}) \tag{5.3}$$

La variable *Alarm* es dependiente de *Burglary* y *Earthquake*, por lo que su probabilidad queda condicionada por la probabilidad conjunta de estas dos variables.

5.1.3.1. Inferencias

Una inferencia es un razonamiento probabilístico que consiste en asignar valores a ciertas variables (evidencia), y se obtiene la probabilidad posterior de las demás variables dadas las variables conocidas.

5.2. Desarrollo

5.2.1. Datasets

Antes de comenzar con el desarrollo del modelo, hubo un inconveniente con respecto a los datasets que se tenían que usar para el entrenamiento.

Los datasets finalmente no fueron proporcionados por los profesionales, por lo que el entrenamiento del clasificador no se podía hacer.

Por lo tanto, a falta de datasets para tomarlos como referencia, se tuvo que usar, como alternativa, datasets generados de forma artificial y automática a través de un script de Python, en lugar de los datasets que supuestamente se debían de haber proporcionado.

Este gran obstáculo trae como consecuencia, que el entrenamiento no sea tan bueno, y que las predicciones sean relativamente inexactas. Esto debido a que los datos generados por el script son aleatorios, por lo que puede dar lugar a datos poco realistas y posibles incoherencias.

5.2.2. Desarrollo del modelo

En un principio, se ha decidido utilizar un clasificador Naive-Bayes para entrenar al modelo.

Toda la información del dataset se extrae y se guarda en un dataframe de la librería *pandas*. Después, se divide los datos en un 75 % para entrenar al modelo y el 25 % restante se usará para testeo.

Internamente, se crea una matriz de confusión, para poder visualizar más fácilmente el número de verdaderos/falsos positivos y negativos.

5.2.3. Entrenamiento del modelo

5.2.4. Medición del rendimiento del modelo

En base al resultado obtenido de esta primera clasificación, se debe evaluar si la clasificación realizada ha sido acertada o no. Para ello, se mide el rendimiento del clasificador mediante el cálculo y la evaluación del p-valor sobre una hipótesis nula.

La hipótesis nula es una afirmación que se supone verdadera mientras no aparezcan evidencias que demuestren lo contrario. En este caso, la hipótesis nula es que el rendimiento del clasificador ha sido de manera puntual y casual.

El p-valor es una medida estadística que indica la probabilidad de que se produzca un resultado aún más alto si la hipótesis nula es cierta.

Por lo tanto, para evaluar el p-valor, se calcula mediante una métrica general el resultado de varias muestras de datos distintas procedentes del mismo conjunto de datos, y posteriormente se comparan con el resultado inicial. Entre las métricas existentes, se pueden aplicar las siguientes:

- **Exactitud (Accuracy)**
- **Precisión**
- **Sensibilidad (Recall)**
- **F1**

Cada métrica devolverá un ratio comprendido entre 0 y 1, y no tienen porqué devolver el mismo resultado. Dependiendo del contexto, el desarrollador es el que tendrá que escoger qué ratio de qué métrica usará para el cálculo del p-valor. Una vez seleccionado la métrica, se calculará el ratio de todos los casos de prueba, y al final se evalúa el nivel de significación. Si dicho nivel es igual o superior al 5 %, entonces se rechaza la hipótesis nula, y, por lo tanto, se puede comprobar que el rendimiento del clasificador no sido de forma casual, sino que es habitual en la mayoría de casos.

5.2.5. Anonimización de los datos

- Técnicas de anonimización de los datos (Marques y Bernardino (2020)).
- Referencia: <https://blog.pangeanic.com/es/6-tecnicas-de-anonimizacion-de-datos-personales/>

Capítulo 6

Conclusiones y Trabajo Futuro

6.1. Conclusiones

En conclusión, el objetivo primordial del proyecto es reducir y prevenir el suicidio en adolescentes, una problemática creciente en estos últimos años. Para ello, se busca poder identificar y predecir con antelación las conductas autolesivas en los jóvenes, para avisar a los profesionales de la salud y que puedan atender a la potencial víctima lo más pronto posible.

Con el desarrollo de este Sistema Experto, se busca cumplir con dicho objetivo mediante la aplicación de la inteligencia artificial, más concretamente, de las redes bayesianas, al campo de la salud, método que se está comenzando a utilizar más en este ámbito. Con el Sistema Experto, los jóvenes y familiares pueden recibir una predicción del modelo, que se hará con gran exactitud ya que estará basado en datos de otros pacientes anteriores, que se habrán usado durante el entrenamiento y el testeo.

6.2. Trabajo a futuro

Con respecto al trabajo a futuro, se ha planteado la posibilidad de llevar el trabajo del Sistema Experto a la nube. Desafortunadamente, la velocidad ha sido uno de los principales inconvenientes a la hora de escoger una librería de Python para desarrollar el Sistema Experto, ya que la gran mayoría necesitaban demasiado tiempo para realizar los procesos de creación, entrenamiento, testeo y predicción del modelo en una máquina local. En algunos casos, ha provocado fallos en la ejecución. Con esta solución, se podría distribuir la carga de trabajo entre varios servidores que tengan mayor capacidad que una máquina local, agilizando el trabajo y obteniendo respuestas más rápidas.

Por otro lado, se espera que en el futuro se pueda materializar el desarrollo de la aplicación móvil, para que los jóvenes puedan acceder más fácilmente a los cuestionarios del sitio web de SIVARIA y puedan obtener resultados más rápidos del Sistema Experto. Además, también podría ser de utilidad para los profesionales de la salud para poder recoger y observar las estadísticas de los resultados obtenidos del Sistema Experto. La idea de la aplicación es que sea gratuita y multiplataforma, es decir, que esté disponible tanto en Android como en IOS, además de su versión de escritorio.

También está abierta la posibilidad de poder mejorar algunos campos del diagrama de la red de Bayes, añadiendo más variables, más tipos de resultados para una variable en concreto o añadiendo más dependencias entre ellas. Todo esto se realizará en concordancia

con los profesionales de la salud involucrados en el proyecto. De esta manera, el modelo puede refinar cada vez más, obteniendo predicciones más exactas.

Capítulo 7

Conclusions and Future Work

7.1. Conclusions

In conclusion, the main objective of the project is to reduce and prevent suicide in adolescents, a growing problem in recent years. To this end, the aim is to be able to identify and predict self-harming behaviour in young people in advance, in order to warn health professionals so that they can attend to the potential victim as soon as possible.

With the development of this Expert System, the aim is to fulfil this objective by applying artificial intelligence, more specifically Bayesian networks, to the field of health, a method that is beginning to be used more in this field. The development of the Expert System will allow young people and relatives receive a highly accurate prediction from the model, based on training data coming from previous patients that had similar pathologies.

7.2. Future work

Regarding the future work, the possibility has been raised of taking the work of the Expert System to the cloud. Unfortunately, the speed is one of the main inconveniences of a model design, since the python libraries generally take too much time to do all the creating, training, testing and predicting processes in a local machine. By this solution, the workload could be distributed among several servers, speeding up the work and obtaining faster responses.

On the other hand, it is expected that in the future the development of the mobile application will materialise, so that young people can more easily access the questionnaires on the SIVARIA website and obtain faster results from the Expert System. get quicker results from the Expert System. In addition, it could also be useful for health professionals to be able to collect and observe the statistics of the results obtained from the Expert System. The idea of the application is that it should be free and multiplatform, i.e. available on both Android and IOS, in addition to its desktop version.

It is also open the possibility of improving some fields of the Bayes network diagram, adding more variables, more types of results for a specific variable or adding more dependencies between them. This will be done in agreement with the health professionals involved in the project. In this way, the model can be further refined, resulting in more accurate predictions.

Bibliografía

- AL-HALABÍ, S. y FONSECA-PEDRERO, E. Suicidal behavior prevention: The time to act is now. *Clínica y Salud*, vol. 32(2), páginas 89–92, 2021.
- ANKAN, A. y PANDA, A. pgmpy: Probabilistic graphical models using python. En *Proceedings of the Python in Science Conference*, SciPy. SciPy, 2015. ISSN 2575-9752.
- ATIENZA, D., BIELZA, C. y LARRAÑAGA, P. Pybnesian: An extensible python package for bayesian networks. *Neurocomputing*, vol. 504, páginas 204–209, 2022.
- ETEROVIC, N. K. H. El observatorio del suicidio, impulsor de la red de prevención del suicidio de las islas baleares. ????
- FONSEKA, T. M., BHAT, V. y KENNEDY, S. H. The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. *Australian & New Zealand Journal of Psychiatry*, vol. 53(10), páginas 954–964, 2019.
- GORE, F. M., BLOEM, P. J., PATTON, G. C., FERGUSON, J., JOSEPH, V., COFFEY, C., SAWYER, S. M. y MATHERS, C. D. Global burden of disease in young people aged 10–24 years: a systematic analysis. *The Lancet*, vol. 377(9783), páginas 2093–2102, 2011.
- MARQUES, J. F. y BERNARDINO, J. Analysis of data anonymization techniques. En *KEOD*, páginas 235–241. 2020.
- MUELA, A., GARCÍA-ORMAZA, J. y SANSINENA, E. Suicidal behavior and deliberate self-harm: A major challenge for youth residential care in spain. *Children and Youth Services Review*, vol. 158, página 107465, 2024. ISSN 0190-7409.
- ORGANIZATION, W. H. ET AL. *Preventing suicide: A global imperative*. World Health Organization, 2014.
- PYPL. Pypl. 2023. Disponible en <https://pypl.github.io/PYPL.html> (último acceso, Abril, 2024).
- PYTHON. *Documentación de Python*. Versión electrónica, 2020. Disponible en <https://docs.python.org/3.9/> (último acceso, Febrero, 2024).
- REACT. Documentación de react native. 2015. Disponible en <https://reactnative.dev/docs/getting-started> (último acceso, Enero, 2024).
- SCHREIBER, J. pomegranate. *Software download*. URL <https://github.com/jmschrei/pomegranate>, 2014.

- SUCAR, L. E. y TONANTZINTLA, M. Redes bayesianas. *Aprendizaje Automático: conceptos básicos y avanzados*, vol. 77, página 100, 2006.
- URBAN, M. *An introduction to LATEX*. TEX users group, 1986.
- WIKIPEDIA. Github — wikipedia, la enciclopedia libre. 2024a. [Internet; descargado 2-abril-2024].
- WIKIPEDIA. React — wikipedia, la enciclopedia libre. 2024b. [Internet; descargado 11-abril-2024].

Apéndice A

Título del Apéndice A

A.1. Instrucciones para ejecutar el script

