

**PONTIFICIA UNIVERSIDAD
CATÓLICA DEL PERÚ**

Escuela de Posgrado



Método semi-supervisado para detectar, clasificar y anotar en un corpus de suicidio textos extraídos de entornos digitales

Tesis para obtener el grado académico de Doctor en Ingeniería que
presenta:

Roberto Wellington Acuña Caicedo

Asesor:

Dr. Héctor Andrés Melgar Sasieta

Lima, 2022

Anexo 2: Modelo de Informe de Similitud


Informe de Similitud

Yo, Andrés Melgar Sasieta, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis titulada titulada MÉTODO SEMI SUPERVISADO PARA DETECTAR, CLASIFICAR Y ANOTAR EN UN CORPUS DE SUICIDIO TEXTOS EXTRAIDOS DE ENTORNOS DIGITALES, del autor ROBERTO WELLINGTON ACUÑA CAICEDO dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 18%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 24/11/2022.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de Suficiencia Profesional, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:

Lima, 28 de noviembre del 2022

Apellidos y nombres del asesor / de la asesora: <u>Melgar Sasieta, Héctor Andrés</u>	
DNI: 10145776	Firma 
ORCID: 0000-0003-1110-123X https://orcid.org/0000-0003-1110-123X	

Resumen

Palabras claves: Informática, Ideación suicida, Clasificadores supervisados, Aprendizaje automático, Redes sociales, Clasificación automática, Suicidio.

La presente tesis doctoral, con un enfoque cualicuantitativo (mixto), se enmarca en la línea del análisis de sentimientos en redes sociales, forma parte del proyecto Life, que busca crear una plataforma integral para detectar y brindar apoyo especializado a usuarios de redes sociales que publican textos con contenido suicida. Por ello se desarrolló el Corpus Life para realizar experimentos con algoritmos de aprendizaje automático, mismo que originalmente constaba de 102 mensajes suicidas (71 textos en inglés y 31 textos en español), 70 de estas muestras Sin Riesgo y 32 con Riesgo. Pero debido al escaso número de muestras y al desbalance entre ellas, los resultados generados no eran confiables.

Por ello esta investigación tuvo como objetivo general desarrollar un método semi-supervisado para detectar, clasificar y anotar en el Corpus Life, textos extraídos de entornos digitales, con el fin de incrementar su número de anotaciones, mediante un proceso de evaluación automática de su calidad, previo a su inclusión o exclusión. Anotaciones que fueron evaluadas manualmente, utilizando para ello la medida de concordancia Cohen's Kappa, con la participación de anotadores especializados quienes evaluaron los textos, alcanzando un nivel de acuerdo entre anotadores de 0,86, cercano al 0,78-0,81 de significancia estadística alcanzado automáticamente por medio del índice macro f1, con el método semi-supervisado. Lo que conllevó a alcanzar experimentos de un mayor grado de confiabilidad, por medio de un método estructurado con actividades, roles y procesos bien definidos y enlazados.

Abstract

Keywords: Computer science, Suicidal ideation, Supervised classifiers, Machine learning, Social networks, Automatic classification, Suicide.

This doctoral thesis with a qualitative-quantitative (mixed) approach is part of the analysis of feelings in social networks that publish texts with suicidal content. For this reason, Corpus life was developed to carry out experiments with machine learning algorithms, which originally consisted of 102 suicide messages (71 texts in English and 31 texts in Spanish), 70 of these samples without risk and 32 with risk. But due to the small number of samples and the imbalance between them, the generated outcome was not reliable.

Therefore, this research had the general objective of developing a semi-supervised method to detect, classify and annotate in the Corpus Life, texts extracted from digital environments, in order to increase their number of annotations, through a process of automatic assessments of their quality, prior to their inclusion or exclusion.

Records which were tested manually, using the Cohen's Kappa concordance measure, with the participation of specialized annotators who evaluated the texts, reaching a level of agreement between annotators of 0.86, close to 0.78-0.81 of statistically significant reaching automatically by means of the f1 macro index, with the semi-supervised method. This led to achieving experiments with a higher degree of reliability, through a structured method with well-defined and linked activities, roles and processes.

Índice

	Pág
Declaración jurada de autenticidad	i
Resumen	ii
Índice de Tablas.....	4
Índice de Figuras.....	6
Capítulo I.....	8
1. Introducción.....	8
1.2. Árbol de problemas.....	11
1.3. Problemática.....	11
1.4. Objetivos.....	14
1.4.1. General.....	14
1.4.2. Específicos.....	14
1.5. Hipótesis.....	14
1.6. Resultados esperados.....	14
1.7. Métodos y procedimientos.....	16
1.8. Metodología de Investigación.....	19
1.9. Alcance del proyecto.....	20
1.9.1. Alcance.....	20
1.9.2. Limitaciones.....	21
1.10. Justificación.....	21
1.11. Viabilidad.....	22
1.11.1. Viabilidad técnica.....	23
1.11.2. Viabilidad económica.....	23
1.11.3. Viabilidad temporal.....	24
Capítulo II.....	26
2. Marco conceptual.....	26
2.1. Suicidio.....	26
2.2. Prevención de suicidio.....	26
2.3. Corpus.....	27
2.3.1. Corpus supervisado.....	28
2.3.2. Corpus semi-supervisado.....	28
2.4. Corpus paralelos.....	29
2.5 Análisis de sentimientos.....	29

Capítulo III.....	30
3. Estado de la Cuestión.....	30
3.1. Definición del Proceso de Revisión Sistemática.....	30
3.1.1. Herramientas de apoyo para la revisión sistemática.	31
3.2. Desarrollo de la revisión sistemática literaria Prevención de Suicidios en Redes Sociales.	32
3.2.1. Objetivos.	32
3.2.2. Definición de preguntas de investigación.....	33
3.2.3. Palabras claves.....	34
3.2.4. Cadena de Búsqueda	34
3.2.5. Criterios de inclusión y exclusión.....	38
3.2.6. Lista de verificación de evaluación de la calidad.....	39
3.2.7. Formulario de extracción de datos	39
3.2.8. Importación de estudios.....	41
3.2.9. Selección de estudios	41
3.2.10. Desarrollo de las preguntas de investigación.....	44
3.3. Desarrollo de metodología de Revisión Sistemática Literaria Análisis de Sentimientos en Redes Sociales.....	50
3.3.1. Preguntas estructuradas PICOC.....	50
3.3.2. Cadena de Búsqueda.	50
3.3.3. Criterios de inclusión y exclusión.....	53
Capítulo IV.....	55
4. Análisis de métodos para anotación de Corpus.....	55
4.1. Evolución de la información	55
4.2. Corpus Lingüísticos	55
Capítulo V.....	58
5. Métodos de evaluación de calidad de anotaciones.	58
5.1. Introducción.....	58
5.2. Pasos del análisis de contenido	59
5.3. Índices estadísticos	59
5.3.3. Selección del Índice estadístico del Corpus Life.....	64
5.3.4. Software de procesamiento de texto.....	65
5.3.5. Esquema de anotación del Corpus Life	66
Capítulo VI.....	73
6. Método de traducción paralela.....	73
6.1. Introducción.....	73

6.2. Proyectos realizados con corpus paralelos.....	73
6.3. Traducción paralela del Corpus de suicidio	74
Capítulo VII.....	76
7. Análisis de resultados y validación de la propuesta.....	76
7.1. Experimentación 1.....	76
7.1.1. Evaluación de clasificadores supervisados para la tarea de detección de mensajes con ideación suicida.....	76
7.1.2. Enfoques de aprendizaje automático	77
7.1.3. Resultados	81
7.1.4. Discusión.....	85
7.2. Experimentación 2.....	86
7.2.1. Anotación semi-supervisada de mensajes potencialmente suicidas en un corpus supervisado	86
7.2.2. Resultados	95
7.2.3. Discusión.....	99
Capítulo VIII	101
8. Conclusiones.....	101
8.1. Conclusiones experimentación 1	101
8.2. Conclusiones experimentación 2	102
8.3. Conclusiones generales	103
Referencias Bibliográficas.....	104

Índice de Tablas.

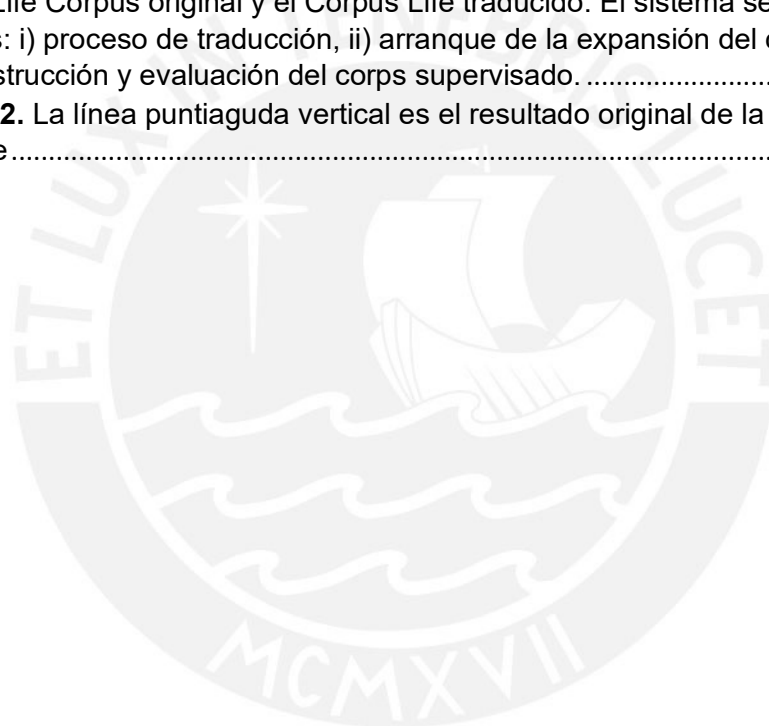
Tabla 1. Medios de verificación de los resultados esperados del primer objetivo específico.	15
Tabla 2. Medios de verificación de los resultados esperados del segundo objetivo específico.	16
Tabla 3. Medios de verificación de los resultados esperados del tercer objetivo específico.	16
Tabla 4. Métodos y Procedimientos primer objetivo específico.	17
Tabla 5. Métodos y Procedimientos segundo objetivo específico.	17
Tabla 6. Métodos y Procedimientos tercer objetivo específico.	17
Tabla 7. Metodología PICOC.	34
Tabla 8. Cadena de búsqueda Web of Science.	35
Tabla 9. Autores más citados Web of Science.	36
Tabla 10. Cadena de búsqueda Scopus.	37
Tabla 11. Autores más citados Scopus.	38
Tabla 12. Criterios de inclusión y exclusión.	39
Tabla 13. Listas de verificación de evaluación de la calidad.	39
Tabla 14. Formulario de extracción de datos.	40
Tabla 15. Artículos encontrados.	41
Tabla 16. Artículos relevantes Scopus.	42
Tabla 17. Artículos relevantes Web of Science.	43
Tabla 18. Artículos relevantes.	43
Tabla 19. Información de los investigadores.	46
Tabla 20. Métodos.	46
Tabla 21. Medidas de evaluación estadísticas.	48
Tabla 22. Algoritmos de clasificación.	49
Tabla 23. Vacíos de conocimiento.	50
Tabla 24. Preguntas estructuradas PICOC.	50
Tabla 25. Cadena de búsqueda Web of Science.	51
Tabla 26. Autores más citados Web of Science.	51
Tabla 27. Cadena de búsqueda Scopus.	52
Tabla 28. Autores más citados Scopus.	53
Tabla 29. Criterios de inclusión y exclusión.	53
Tabla 30. Numero de ejemplos por cada tipo de alerta.	71
Tabla 31. Numero de ejemplos por tipo de alerta riesgo y No riesgo.	72
Tabla 32. Textos en idioma inglés, por nivel de alerta.	75
Tabla 33. Numero de características por cada tipo de característica, con y sin stopwords.	77
Tabla 34. Algoritmos y parámetros por defectos utilizados in los experimentos.	79
Tabla 35. Resultados de los principales valores de la evaluación del Corpus con diferentes características y diferentes clasificadores para la versión del Corpus con 4 clases.	82
Tabla 36. Resultados de los principales valores de la evaluación con diferente clasificador y diferentes categorías para la versión del Corpus con 2 clases (“Riesgo” y “No Riesgo”).	83
Tabla 37. Los clasificadores y características utilizados para los modelos de aprendizaje más relevantes siguen el PBI.	85

Tabla 38. Número de muestras para cada tipo de "Nivel de alerta".....	87
Tabla 39. Número de muestras para cada tipo de "Nivel de alerta".....	87
Tabla 40. Numero de Ejemplos y limites por iteración.....	89
Tabla 41. Acuerdos entre revisores.....	90
Tabla 42. Corpus utilizado en experimentos.....	90
Tabla 43. Macro F1, Macro Precision, y Macro Recall. Corpus en idioma inglés combinado con características de entrenamiento. El intervalo de confianza se calculó con $p < 0,01$	97



Índice de Figuras.

Ilustración 1 Usuarios de servicios de comunicación, a febrero del 2022.....	8
Ilustración 2 Ciclo del Proyecto Life.....	21
Ilustración 3 Categorías de Web of Science.....	35
Ilustración 4 Red bibliométrica de palabras claves Web of Science.....	36
Ilustración 5 Categorías de Scopus.....	37
Ilustración 6 Red bibliométrica de palabras claves Scopus.	38
Ilustración 7 Red bibliométrica de palabras claves Scopus.	52
Ilustración 8 Red bibliométrica de palabras claves Scopus.	53
Ilustración 9. Método semi-supervisado.....	57
Ilustración 10 Representación del Índice Bidimensional de Precisión (PBI) de los resultados del Corpus con dos clases.....	84
Ilustración 11. Esquema de flujo de trabajo del sistema. El sistema se evaluó utilizando el Life Corpus original y el Corpus Life traducido. El sistema se compone de tres procesos: i) proceso de traducción, ii) arranque de la expansión del corpus, iii) revisión, construcción y evaluación del corpus supervisado.	88
Ilustración 12. La línea puntiaguda vertical es el resultado original de la medida f con el Corpus Life.....	96



Índice de Formulas.

Ecuación 1. Cohen's kappa	56
Ecuación 2. Percent agreement	60
Ecuación 3. F-measure	78
Ecuación 4. ROC Area	78
Ecuación 5. Precision	78
Ecuación 6. Recall	78
Ecuación 7. Precision_risk.....	80
Ecuación 8. Precision_No_risk.....	80
Ecuación 9. Estadístico Chi2,.....	80
Ecuación 10. SS_total	81
Ecuación 11. SS_trat.....	81
Ecuación 12. SS_error	81
Ecuación 13. F.....	81
Ecuación 14. Accuracy.....	92
Ecuación 15. Balanced_accuracy.....	92
Ecuación 16. Micro F1-score.....	93
Ecuación 17. Macro F1-score.....	93
Ecuación 18. Micro_precision.....	93
Ecuación 19. Macro_precision	93
Ecuación 20. Micro_recall.....	94
Ecuación 21. Macro_recall.....	94
Ecuación 22. Jaccard similarity	94

Capítulo I

1. Introducción

A febrero del año 2022, el planeta cuenta con una población aproximada de 7.925 millones de personas [1]. De las cuales 5.310 millones son usuarios únicos de telefonía móvil (Penetración de mercado 67,1%), 4.950 millones son usuarios de Internet (Penetración de mercado 62,5%), y 4.620 millones son usuarios activos de redes sociales (Penetración de mercado 58,4%). Servicios que se han desarrollado dada la disrupción tecnológica que se ha dado en las últimas décadas [2]. Ilustración 1.

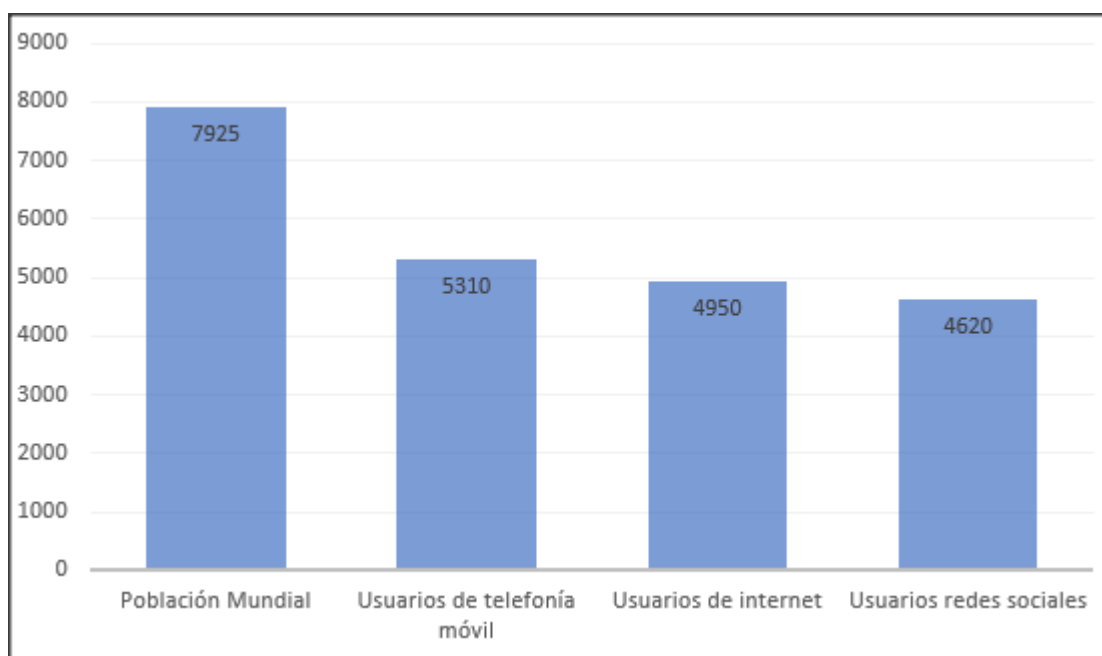


Ilustración 1 Usuarios de servicios de comunicación, a febrero del 2022

Hoy en día, las Tecnologías de la Información y Comunicación (TIC), influyen en el desarrollo de la cultura y la sociedad dada la forma en que los individuos interactúan, interpreten información y se comunican entre sí por medio de las redes sociales [3], [4][5][6][7]. En las cuales consumen y crean contenido a través de la televisión digital, interacción en redes sociales, envío de imágenes y videos, etc. Que ha llevado, a que la humanidad genere aproximadamente 44 zettabytes¹ de datos digitales hasta el año 2020, y se estima que llegara a 175 zettabytes hasta el año 2025 [8].

De este universo digital de información, el 33% está constituida por datos con etiquetas, los cuales contienen información que describen al dato, mientras que el 67% no cuentan con estas etiquetas [8]. Analizar parte del gran volumen de información existente, representa un gran desafío y una oportunidad para especialistas en las diferentes áreas del conocimiento [9]. A través del aprendizaje automático o Machine Learning, el cual es un método analítico que permite que un sistema por sí mismo sin intervención humana y en forma automatizada, aprenda a descubrir patrones, tendencias y

¹ Unidad de almacenamiento de información equivalente a 10^{21} .

relaciones entre datos. El aprendizaje automático permite a las computadoras aprender de forma supervisada [10], no supervisado [11][12], o semi-supervisada [13].

El aprendizaje supervisado es una técnica que induce una función a partir de un conjunto de datos etiquetados previamente, conocidos como datos de entrenamiento los cuales consisten en pares de objetos tanto de entrada como de salida. El cual puede resultar costoso computacionalmente, y con un alto consumo de tiempo, tanto en el proceso de anotación, como de evaluación [14]. En cambio, el aprendizaje semi-supervisado, es una técnica de aprendizaje automática que utiliza datos de entrenamiento tanto etiquetados como no etiquetados, para comprender como la combinación de ambos puede generar nuevos aprendizajes, con un costo computacional menor [13].

Aprendizajes, con los que se puede analizar contenido útil, compartido por usuarios de redes sociales, acerca de situaciones de la vida real, como el marketing [15], cursos en línea [16], opinión de productos [17], ubicación geográfica, preferencias alimenticias [18], pasatiempos, tienda favorita, política [19], identificación de personas socialmente peligrosas [20], suicidas [21][22][23][24][25][26][27][28][29], e influencers, quienes influyen de manera positiva o negativa en las personas [30]. En el contexto negativo pueden influir contagios masivos entre sujetos vulnerables que experimentan angustia o trastorno mental, que son algunos de los predictores más fuertes de conductas suicidas [31][32][33].

Los algoritmos de aprendizaje automático han podido de manera exitosa identificar notas suicidas provenientes de grupos de noticias y medios sociales [34][35][36]. Distinguir entre notas escritas por personas que murieron por suicidio y notas de suicidio simuladas, mejor que los profesionales de salud mental (71% vs. 70%) [37]. Además, a partir de notas de suicidio, y datos de microblogs, han posibilitado la construcción de modelos de aprendizaje automático, para identificar usuarios con sentimientos suicidas con aproximadamente un 90% de precisión [27].

En el año 2022, la cifra de muertes por suicidio será de aproximadamente 911.375 y el número de intentos suicidas fallidos de 18.227.500. De acuerdo con la tasa mundial anual estandarizada de suicidios de 11,4 muertes por cada 100.000 habitantes y 20 intentos fallidos por cada muerte [38]. Hay que tener en cuenta que estos números pueden ser mayores, ya que, en países con buenos datos de registro civil, muchas veces se clasifican estas muertes como accidentales y en países sin registros fiables, los suicidios ni se cuentan [38].

Las muertes por suicidio se dan tanto en países ricos de acuerdo a su producto interno bruto (25% aproximadamente del total mundial de suicidios), como en países pobres y medianos (75% aproximadamente del total mundial de suicidios) [39]. El suicidio afecta a todos los grupos de edad, siendo más alto en personas de 70 años o más [40][41][42], en el año 2015 fue la segunda causa de muerte violenta en el grupo de edad de 15 a 44 años [39][43]. El suicidio afecta tanto al suicida, como a sus familiares y allegados [44][45].

En el año 2013, en el marco de la 66ª Asamblea Mundial de la Salud, los estados miembros de la OMS entre los cuales se encuentran Perú, España y Ecuador se comprometieron entre otros temas a reducir las tasas nacionales de suicidios en un 10% para el año 2020 [38]. Para tal efecto acordaron elaborar y poner en práctica estrategias nacionales integrales de prevención de suicidios, reforzando para ello sus sistemas de información, datos científicos, investigación y colaboración universitaria sobre salud mental, prestando especial atención a los colectivos de mayor riesgo suicida, como son:

homosexuales, lesbianas, bisexuales, transexuales, jóvenes, refugiados, migrantes u cualquier otro grupo vulnerable [38][39][45][46].

Por lo general, un suicida pasa por un período de profundo sufrimiento personal con frecuencia en silencio, antes de tomar la fatídica decisión de terminar con su propia vida. Por lo que predecir que alguien se va a suicidar ha sido una tarea imposible de determinar [47][48][49][50][51]. Pero si es posible detectar factores de riesgo suicida, a través de herramientas clínicas estándar operadas por médicos bien entrenados [52][53]. Mientras que la informática, específicamente el Procesamiento del Lenguaje Natural (NLP), ofrece la oportunidad de comprender los indicadores de pensamientos suicidas a través de la interacción entre las computadoras y el lenguaje humano [54][55], cuando se expresan en forma escrita y hablada [56][57][58][59]. Permitiendo comprender de una mejor manera sus pensamientos, sentimientos, creencias, comportamientos y personalidades [60], con los cuales es posible analizar y predecir resultados [61][62].

El éxito del análisis de sentimientos [63][64][65][66], depende en gran medida de la calidad de los datos implícitos [67] y no específicos [68] que se pueden extraer del gran flujo de información ahora disponible [69][70], en las redes sociales [63]. Con los datos recopilados, compilados, y etiquetados de forma manual o automática a partir de textos orales o escritos, se puede crear un Corpus [71][72] sea este supervisado [10][73], o semi-supervisado [10][11]. Un Corpus es un conjunto estructurado de ejemplos reales orales o escritos de un tema específico, creado de forma independiente a su uso [74].

En el proceso de creación de un Corpus, una vez ubicados y recopilados los textos desde redes sociales u otras fuentes, estos deben ser procesados de preferencia de forma supervisada. Lo que implica, etiquetar manualmente los datos y determinar su calidad a través de acuerdos entre anotadores con la ayuda de métodos estadísticos como el coeficiente de Cohen's [75][76][77][78][25][79][80][81], realizado por grupos específicos de expertos [82][83][84] o por múltiples grupos de expertos a través de trabajo colaborativo [84][85]. Solo los textos que logran alcanzar un alto nivel de acuerdo serán registrados en el Corpus.

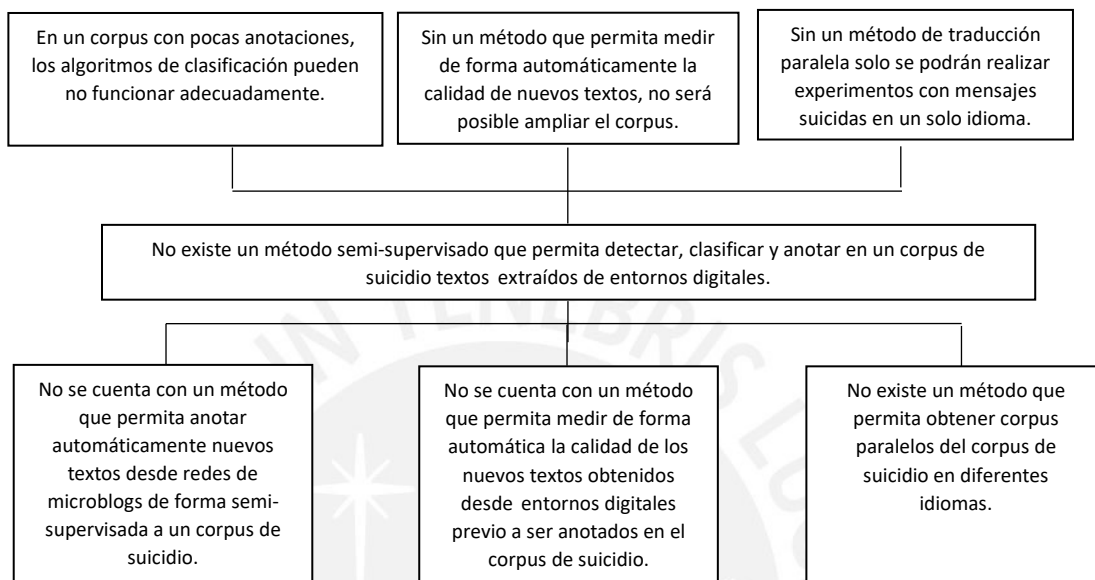
Los textos que los diferentes grupos de investigación ubican en las redes sociales para el desarrollo de sus Corpus, generalmente se centran en una sola fuente de datos como: Twitter [86][87][25][88][89], Weibo [90][91], Netlog [92], Reddit [93], u otros microblogs [94], igualmente hay experimentos con dos fuentes de datos [85] y otros con más de dos fuentes [95].

A los Corpus anotados de forma supervisada que alcanzan cierta calidad medida objetivamente, por lo general se los conocen como Corpus estándar de oro [96][97][84][98], los cuales son utilizados conjuntamente con algoritmos de clasificación, para entrenar y agregar nuevos datos en muchos casos de forma semi-supervisada [99][100][101][87][102][25][103][104]. Entre las diferentes áreas de conocimiento en las cuales grupos de investigación han desarrollado Corpus tenemos: suicidio [95], abuso de medicamentos [102], depresión [87], especies filipinas [105], acoso cibernético [103], agricultura [99], etc.

La presente investigación busca desarrollar un método semi-supervisado que detecte, clasifique y anote en un Corpus de suicidio textos extraídos de entornos digitales, con el fin de incrementar el número de anotaciones de corpus de ideación suicida a partir de textos de redes sociales. Por medio de establecer un proceso de evaluación que permita

de forma automática medir la calidad de los textos, previo a su inclusión o exclusión en Corpus estándar de oro que puedan ser Corpus paralelos, es decir, que puedan utilizarse en diferentes idiomas.

1.2. Árbol de problemas



1.3. Problemática

En el año 2013, la Organización Mundial de la Salud (OMS) reconoció al suicidio como una prioridad de salud pública [39]. Con una población mundial al año 2022, de más de 7.925 millones de habitantes [1], y una tasa internacional anual estandarizada de 11,4 suicidios por cada 100.000 habitantes [106].

En los países ricos según su PIB per cápita, ocurren aproximadamente el 25% del total mundial de muertes por suicidios, en los cuales, por cada mujer que se suicida, se suicidan 3 hombres. Mientras que en los países pobres y medianos el porcentaje de suicidios llega aproximadamente al 75%, en los cuales por cada mujer que se suicida, se suicidan 1,5 hombres [39][107]. En el año 2015, el suicidio fue la segunda causa principal de muerte violenta en el mundo, para las personas en el rango de edad entre 15 a 44 años, y en Europa la primera [108]. En el año 2017, el 17,2% de los estudiantes de secundaria de EEUU informó haber considerado seriamente intentar suicidarse y el 7,4% informó haberlo intentado el año anterior [109]. En el año 2018 por causas del suicidio fallecieron en el Perú aproximadamente 1.912 personas, en el Ecuador 1.294 y en España 3.569 [110].

En 1959, se creó el primer conjunto ordenado de notas originales de suicidio, conocido como Corpus Genuine Suicide Notes (GSN), que identificaba características textuales de notas suicidas, dando inicio a la investigación sobre sistemas automáticos para la detección de suicidios en mensajes escritos [111]. El Corpus GSN contenía una muestra de 66 notas de suicidio, de las cuales la mitad eran auténticas y la otra mitad simuladas

[111]. Sin embargo, estos trabajos no son aptos para la detección de mensajes suicidas, sino solo para la diferenciación entre notas suicidas reales y falsas, ya que estos Corpus no tienen ejemplo de mensajes suicidas, ni otro tipo de mensajes que no sea notas suicidas [112][113][114][115][116][117][118].

Actualmente, las redes sociales son útiles para estudiar conductas de personas suicidas, por parte de grupos de investigación adscritos a universidades, sociedad, gobiernos y entidades de salud internacional [119]. Para entre otras prioridades, intentar dar cumplimiento al compromiso de sus países con la OMS, de reducir las tasas de suicidio en un 10% hasta el año 2020 [39].

Pero, por no contar con un Corpus supervisado o Gold Standard de acceso abierto orientado al suicidio, no es posible entrenar nuevos textos obtenidos desde redes de microblogs, blogs, foros, notas suicidas, etc. Por ello es imperativo su desarrollo con el fin de realizar experimentos en función de detectar mensajes con ideación suicida en redes sociales [71].

Para que un Corpus de mensajes suicidas, generado a partir de textos ubicados y recopilados desde redes sociales u otras fuentes, pueda convertirse en un referente para los grupos que investigan en esta línea del conocimiento, deberá ser desarrollado en base a los siguientes métodos:

Un método que permita detectar, clasificar y anotar automáticamente de forma semi-supervisada nuevos textos relacionados con el suicidio desde redes de microblogs, blogs, foros u otras fuentes. Las cuales están constantemente alimentadas con la publicación de nuevos mensajes tanto escritos como hablados, por parte de los usuarios, en los que expresan sus sentimientos [56][57][58][59]. Con el objeto de incrementar constantemente el número de muestras del Corpus, para poder realizar búsquedas cada vez más exactas de potenciales usuarios suicidas en redes sociales.

Un gran número de muestras de discursos naturales permite la realización de análisis empíricos de patrones reales de uso en un idioma, que al combinarse con herramientas computacionales hace posible la realización de análisis de un alcance que de otro modo no sería factible por los grandes volúmenes de datos que se deben procesar [120]. Por el contrario, los experimentos que se realicen con un Corpus con pocas anotaciones no siempre darán los mejores resultados, por no contar con suficientes ejemplos que les permita desarrollar patrones reales de uso lingüístico.

Este método de incremento de muestras al Corpus deberá de forma semi-supervisada, procesar nuevos mensajes y de forma automática anotar aquellas muestras validadas por algoritmos de Machine Learning, que basen su entrenamiento en las muestras iniciales anotadas de forma manual y cuya calidad haya sido medida por acuerdos.

Un método que permita medir de forma automática la calidad de los nuevos textos obtenidos desde entornos digitales. Para alcanzar resultados confiables en los diferentes experimentos que los grupos de investigación realicen utilizando el Corpus de suicidio, la calidad de sus muestras o ejemplos, debe ser medida. Por ello en su fase inicial, debe desarrollarse de forma

supervisada, en el cual la calidad de sus anotaciones debe ser medida mediante acuerdos, utilizando de preferencia el estadístico Cohen's Kappa [121]. Cuya metodología, tiene en cuenta la concordancia al azar entre evaluadores expertos [122][123][124], respecto a la calidad de las muestras analizadas, para determina su inclusión o exclusión.

Conduciendo a clasificaciones con un mayor grado de precisión, y soluciones de subconjuntos de características relevantes [4][107]. Garantizando que las anotaciones del Corpus sean de calidad, siendo por esto, ampliamente utilizado en proyectos provenientes del ámbito de redes sociales, ciencias, biología, ciencias médicas, entre otros, teniendo como desventaja los altos costos humanos e informáticos que representa su utilización [5].

Por lo tanto, las siguientes anotaciones del Corpus, deberán ser procesadas de formas semi-supervisada, basando el entrenamiento de los algoritmos de *Machine Learning* en las muestras supervisadas. Con un procesamiento semi supervisado de nuevas muestras, y una selección automática previo a su anotación al Corpus, los costos humanos e informáticos disminuirían, respecto a que si el proceso de incremento de nuevas muestras se lo realizara de forma supervisada [10].

Un método que permita obtener Corpus paralelos en diferentes idiomas, a partir del Corpus de suicidio. Proceso necesario, dado que actualmente en el mundo existen alrededor de 7.139 idiomas, de entre los cuales los más hablados son el inglés, chino mandarín, hindú, español, árabe, bengalí, francés, ruso, portugués y urdu [125].

Por esto el Corpus de suicidio, no solo deberá mantener sus muestras en idioma inglés, sino que deberá buscar métodos, por medio de los cuales dada la necesidad generar procesos de traducción paralela a los diferentes idiomas que se requieran [73].

Los Corpus paralelos utilizan un método que permite realizar una traducción automática, desde el texto fuente al texto destino, utilizando entre otros la alineación de palabras como puente [126] [127][128].

Posibilitando de esta manera, que los diferentes grupos de investigación en el campo de la prevención de suicidios puedan realizar sus respectivos experimentos en su idioma nativo, convirtiendo a cada uno de estos Corpus en una herramienta que posibilite en diferentes idiomas ubicar posibles suicidas en redes sociales, con el fin de poderles ayudar por medio de personal especializado.

Además, este Corpus en su diseño deberá estar estructurado por los niveles de alerta **Riesgo** y **No Riesgo**, los cuales deberán en lo posible estar balanceados con un número similar de muestras. Con el fin de alcanzar resultados óptimos al momento de realizar el proceso de búsqueda de usuarios suicidas en redes sociales, así como el entrenamiento de nuevos textos con algoritmos de clasificación.

1.4. Objetivos

1.4.1. General

Desarrollar un método semi-supervisado que detecte clasifique y anote en un Corpus de suicidio textos extraídos de entornos digitales

1.4.2. Específicos

Proponer un método semi-supervisado que detecte, clasifique y anote en un Corpus de suicidio textos extraídos de entornos digitales.

Realizar la validación de la calidad de los textos relacionados con el suicidio obtenidos desde entornos digitales.

Generar Corpus paralelos a diferentes idiomas a partir del Corpus de suicidio escrito en idioma inglés.

1.5. Hipótesis

La formulación de la hipótesis es este proyecto de tesis se centra en demostrar la viabilidad de desarrollar un método que permita anotar automáticamente nuevos textos desde redes de microblogs de forma semi-supervisada a un Corpus de suicidio. Para lograr este objetivo es necesario validar y contrastar esta propuesta con trabajos de investigación similares que pudieren existir.

H_i Si se desarrolla un método semi-supervisado que detecte, clasifique y anote textos relacionados con el suicidio provenientes de entornos digitales entonces se incrementa el número de registros del Corpus de suicidio, lo que permitirá con algoritmos de clasificación entrenar nuevos datos para realizar búsquedas de usuarios suicidas en redes sociales.

1.6. Resultados esperados

Objetivo 1. Proponer un método semi-supervisado que detecte, clasifique y anote en un Corpus de suicidio textos extraídos de entornos digitales.		
Resultados	Meta	Medios de Verificación
(R 1.1) Documentos de análisis de técnicas y métodos reportados utilizados en la literatura	Documento de Tesis	Documento que describe el método semi supervisado utilizado para incrementar anotaciones

respecto a la detección, clasificación y anotación automática de textos desde redes de microblogs de forma semi-supervisada a un Corpus de suicidio.		automáticamente en el Corpus de suicidio. Capítulo VI.
	Corpus	Corpus de suicidio con 273 registros.
(R 1.2) Formulario de extracción de datos, realizado en base a las preguntas de investigación de la revisión sistemática literaria.	Formularios	Formularios de extracción de datos que permiten extraer información de cada artículo que haya sido incluido en esta investigación. Sub Capítulo 3.2.7.
(R 1.3) Análisis comparativos de la información recopilada en los formularios de extracción de datos de los artículos incluidos en esta investigación.	Documento de Tesis	Documento que presenta los resultados estadísticos de los datos recopilados por medio del formulario de extracción de conocimientos. Capítulo III.
(R 1.4) Documento que define las actividades seguidas en los experimentos realizados, para determinar las mejores características del Corpus y el algoritmo de clasificación idóneo para el proceso de entrenamiento.	Artículos Científicos	Artículos científicos que describen los experimentos realizados en esta investigación para determinar las mejores características del Corpus de suicidio y el algoritmo idóneo para entrenar nuevos datos.

Tabla 1. Medios de verificación de los resultados esperados del primer objetivo específico.

Objetivo 2. Realizar la validación de la calidad de los textos relacionados con el suicidio obtenidos desde entornos digitales.		
Resultados	Meta	Medios de Verificación
(R 2.1) Documentos de análisis de técnicas y métodos reportados utilizados en la literatura que describe el proceso metodológico que permite medir la calidad de los textos de forma automática previo a ser anotados en el Corpus de suicidio.	Documento de Tesis	Documento que describe los métodos estadísticos utilizados actualmente para medir la calidad de textos. Capítulo IV.
	Reporte estadístico	Documento en el cual se presentan los resultados de la evaluación automática de la calidad realizada a 200 textos provenientes de redes sociales.
(R 2.2) Documento que describe las características del Corpus de suicidio, en base a las cuales se debe ubicar cada uno de los textos evaluados, sea esto	Documento de Tesis	Documento de estructura del corpus de Suicidio. Sub Capítulo 6.2.1.

de forma automática o manual.		
-------------------------------	--	--

Tabla 2. Medios de verificación de los resultados esperados del segundo objetivo específico.

Objetivo 3. Generar Corpus paralelos a diferentes idiomas a partir del Corpus de suicidio escrito en idioma inglés.		
Resultados	Meta	Medios de Verificación
(R 2.1) Documento que describe procesos metodológicos que permiten generar Corpus paralelos en diferentes idiomas, a partir de un Corpus principal.	Documento de Tesis	Documento que describe el método de traducción utilizado que permite obtener Corpus paralelos del Corpus de suicidio en diferentes idiomas. Capítulo V.
	Corpus	Corpus de suicidio en idioma español, traducido de forma paralela al principal desarrollado en idioma inglés.

Tabla 3. Medios de verificación de los resultados esperados del tercer objetivo específico.

1.7. Métodos y procedimientos

Objetivo 1. Proponer un método semi-supervisado que detecte, clasifique y anote en un Corpus de suicidio textos extraídos de entornos digitales.	
Resultados	Herramientas, métodos y metodologías.
(R 1.1) Documentos de análisis de técnicas y métodos reportados utilizados en la literatura respecto a la detección, clasificación y anotación automática de textos desde redes de microblogs de forma semi-supervisada a un Corpus de suicidio.	Base de datos bibliográfica Scopus. Base de datos bibliográfica Web of Science. Herramienta en línea de revisión sistemática Parsifal. Software de Desarrollo Python. Software Weka. Librerías de traducción. Microsoft Word Microsoft Excel Estadístico Cohen's Kappa. Estadísticos media, varianza. Prueba T student. Revisión sistemática literaria.
(R 1.2) Formulario de extracción de datos, realizado en base a las preguntas de investigación de la revisión sistemática literaria.	
(R 1.3) Análisis comparativos de la información recopilada en los formularios de extracción de datos de los artículos incluidos en esta investigación.	
(R 1.4) Documento que define las actividades seguidas en los experimentos realizados, para determinar las mejores características del Corpus y el algoritmo de clasificación	

idóneo para el proceso de entrenamiento.	
--	--

Tabla 4. Métodos y Procedimientos primer objetivo específico.

Objetivo 2. Realizar la validación de la calidad de los textos relacionados con el suicidio obtenidos desde entornos digitales.	
Resultados	Herramientas, métodos y metodologías.
(R 2.1) Documentos de análisis de las técnicas y métodos reportados y utilizados en la literatura que describe el proceso metodológico que permite medir la calidad de los textos de forma automática previo a ser anotados en el Corpus de suicidio.	Base de datos bibliográfica Scopus. Base de datos bibliográfica Web of Science. Herramienta en línea de revisión sistemática Parsifal. Microsoft Word Microsoft Excel
(R 2.2) Documento que describe las características del Corpus de suicidio, en base a las cuales se debe ubicar cada uno de los textos evaluados, sea esto de forma automática o manual.	Software de Desarrollo Python. Estadístico Cohen's Kappa. Revisión sistemática literaria.

Tabla 5. Métodos y Procedimientos segundo objetivo específico.

Objetivo 3. Generar Corpus paralelos a diferentes idiomas a partir del Corpus de suicidio escrito en idioma inglés.	
Resultados	Herramientas, métodos y metodologías.
(R 3.1) Documento que describe procesos metodológicos que permiten generar Corpus paralelos en diferentes idiomas, a partir de un Corpus principal.	Base de datos bibliográfica Scopus. Base de datos bibliográfica Web of Science. Herramienta en línea de revisión sistemática Parsifal. Software de Desarrollo Python. Librerías de traducción. Microsoft Word. Microsoft Excel. Estadístico Cohen's Kappa. Revisión sistemática literaria.

Tabla 6. Métodos y Procedimientos tercer objetivo específico.

Revisión sistemática literaria.

Una revisión sistemática de la literatura es un estudio secundario que tiene como objeto identificar, analizar e interpretar todas las evidencias disponibles de estudios primarios relacionados con una pregunta de investigación específica, lo cual implica planificar, realizar e informar la revisión [129].

Parsifal.

Parsifal² es una herramienta en línea diseñada para ayudar a los investigadores a realizar revisiones sistemáticas de la literatura en el contexto de la ingeniería del software. Esta herramienta permite a los investigadores sin importar su ubicación geográfica trabajar juntos en un espacio de trabajo compartido, diseñar el protocolo y realizar la investigación [129].

Base de datos bibliográfica Scopus.

Se trata de una base de datos multidisciplinaria orientada hacia la investigación. Es la mayor base de datos referencial de citas y resúmenes de literatura académica revisada por pares de Elsevier. Además, ofrece herramientas de soporte al proceso de investigación, permite a los usuarios realizar un seguimiento y análisis de revistas, autores y afiliaciones institucionales. Entre sus prestaciones se encuentran los índices bibliométricos **Scimago Journal Rank**³ (SJR), **Source-normalized impact Paper**⁴ (SNIP) y **CiteScore**⁵. Puede utilizarse como herramienta de investigación y para la difusión de los resultados de I+D+I [130].

Base de datos bibliográfica Web of Science.

Web of Science⁶ (WOS), propiedad de Clarivate Analytics, es una colección de bases de datos bibliográficas y citas de artículos científicos, libros y otro tipo de documento revisados por pares. Soporta más de 250 disciplinas, entre las que constan ciencias, ciencias sociales, arte y humanidades. Permite el acceso a las siguientes herramientas **Web of Science**, **SciELO Citation Index**, **Medline**, **Derwent Innovations Index** (1980-2009), **Journal Citation Reports** (JCR), **Essential Science Indicators** (ESI), **Endnote Web**, **ResearcherID**, **KCI Korean Journal Database** y **Russian Science Citation Index** [131].

Python.

Python es un lenguaje de programación orientado a objetos cuyo modelo de datos también está basado en objetos y tiene como filosofía la legibilidad de su código. Se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa, y en menor medida, programación funcional [132].

Función googletrans.

Googletrans es una biblioteca de Python gratuita e ilimitada que implementó la API de Google Translate. Esto utiliza la API Ajax de Google Translate para realizar llamadas a métodos para detectar y traducir textos [133].

² <https://parsif.al>

³ <https://www.scimagojr.com/>

⁴ <http://www.journalindicators.com/>

⁵ <https://www-scopus-com.consultaremota.upb.edu.co/sources>

⁶ <https://clarivate.com/products/web-of-science/>

1.8. Metodología de Investigación

El presente proyecto de tesis doctoral, ubicado en la línea de análisis de sentimientos en redes sociales presenta un enfoque cualicuantitativo (mixto). Los métodos mixtos o híbridos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos tanto cuantitativos como cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada y lograr un mayor entendimiento del fenómeno bajo estudio [134].

En la ruta mixta se utiliza evidencia de datos numéricos, verbales, textuales, visuales, simbólicos y de otras clases para entender problemas en las ciencias, en ese sentido el enfoque cuantitativo, permite generalizar resultados a través de un diseño experimental con un patrón predecible y estructurado. Los estudios previos ubicados por medio de la revisión sistemática literaria permiten generar el modelo teórico a ser comprobado a través de las hipótesis, se utiliza la lógica o razonamiento deductivo, para diseñar un método semi-supervisado que detecte, clasifique y anote en un corpus de suicidio textos extraídos de entornos digitales, en base a un análisis exhaustivo de propuestas desarrolladas en este campo del conocimiento. Finalmente, los datos serán procesados mediante métodos estadísticos para establecer conclusiones válidas, las cuales no deberán ser afectadas por las tendencias del investigador u otras personas.

Con el enfoque cualitativo se garantiza la necesaria evaluación de las anotaciones realizadas de forma automática al corpus de suicidio, mediante una medida de concordancia como Cohen's Kappa, con la participación de anotadores especializados quienes deberán comprender y evaluar los textos de forma manual. En el desarrollo de este enfoque se utiliza una lógica y proceso inductivo que utiliza premisas y observaciones para sacar conclusiones probables.

Las fases que han sido consideradas para la ejecución de la presente investigación se detallan a continuación:

La primera fase del proyecto comprende una revisión sistemática literaria que posibilite la revisión de aspectos cuantitativos y cualitativos de estudios primarios relacionados con la prevención de suicidios en redes sociales. A partir de una cadena de búsqueda en las bases de datos bibliográficas Scopus, Web of Science y Google Scholar. Seguidamente los metadatos de los artículos resultantes de las búsquedas tanto en Scopus como en Web of Science serán exportados a la herramienta en línea Parsifal. En el desarrollo de la revisión sistemática, los textos hallados en las bases de datos serán incluidos o excluidos de la investigación de acuerdo con los criterios de inclusión y exclusión planteados. Los textos incluidos en la investigación serán analizados y sus datos compilados en un formulario de extracción de datos elaborado en base a las preguntas de investigación planteadas. Este proceso tendrá como objeto recopilar la información existente, para alcanzar un alto conocimiento en el área de prevención de suicidios en redes sociales, y a partir de este umbral iniciar el proyecto de tesis doctoral.

La segunda fase contempla la realización de experimentos utilizando técnicas de aprendizaje automático para llevar a cabo un análisis sistemático de todas las combinaciones posibles de características textuales del Corpus Life, con algoritmos de clasificación supervisados con parámetros predeterminados, disponibles en el

software para el aprendizaje automático y la minería de datos Weka. Para a continuación medir la significancia de los resultados estadísticos mediante el análisis de varianza con un factor (ANOVA) y poder determinar el clasificador y características con los mejores resultados. En un segundo grupo de experimentos se diseñará un método que permita anotar de forma semi-supervisada registros automáticos a partir de textos extraídos de plataformas sociales como Reedit (subreddit "SuicideWatch"), Twitter, Weibo, Netlog, entre otros, al Corpus Life, teniendo en cuenta el balance de sus clases.

En la fase tres, de este proyecto se evaluarán las anotaciones realizadas de forma automática al corpus de suicidio, mediante una medida de concordancia como Cohen's Kappa, con la participación de anotadores especializados quienes deberán comprender y evaluar los textos de forma manual y decidir si estos son suicidas o no suicidas, resultado que serán comparados, con los resultados obtenidos mediante el método computacional y con ello poder determinar el nivel de acuerdo.

En la fase cuatro, se procederá a utilizar el método más acorde, para realizar la traducción de textos del Corpus Life a diferentes idiomas.

Finalmente, se obtienen las conclusiones y recomendaciones.

1.9. Alcance del proyecto

El alcance y limitaciones de esta investigación son las siguientes:

1.9.1. Alcance

El presente proyecto de tesis doctoral orienta su estudio al desarrollo de un método semi-supervisado para detectar, clasificar y anotar en un corpus de suicidio textos extraídos de entornos digitales, con el fin de detectar a usuarios de redes sociales que publican intenciones suicidas. Ello con el propósito de cubrir el vacío existente con respecto al desarrollo de metodologías semiautomáticas que permitan incrementar el tamaño del Corpus de ideación suicida, a partir de textos extraídos de redes sociales, para detectar en línea a potenciales usuarios suicidas y que entidades especializadas puedan contactarlos y animarlos a no suicidarse.

La investigación tiene un alcance explicativo transitando por lo exploratorio y descriptivo, pues se enfoca al estudio de fenómenos que no se han investigado previamente con la suficiente sistematicidad. Se tiene el interés de examinar las características del fenómeno, plantear una hipótesis que busca realizar estudios de tipo fenomenológicos o narrativos constructivistas. Para describir las representaciones subjetivas que emergen en un grupo humano sobre un determinado fenómeno, en consecuencia, se busca una explicación y determinación de los fenómenos al aplicar un estudio de tipo predictivo, se establece una relación causal entre al menos dos variables, y se propone una teoría que busca una comprensión de un fenómeno. Por otro lado, genera una manipulación intencionada de la variable independiente que permite comprobar hipótesis que explican el comportamiento del tema estudiado [135].

Con el desarrollo de esta investigación, que es parte del proyecto de prevención de suicidio Life, se pretende aportar a la reducción de las tasas de suicidios que a nivel global tienen un estándar de 11.4 muertes por cada 100.000 habitantes.

1.9.2. Limitaciones

Esta investigación forma parte del proyecto Life, el cual está compuesto de varias etapas tales como: 1. Creación de un Corpus supervisado; 2. modelos computacionales de aprendizaje automático; 3. definición de procesos de anotación; 4. desarrollo de sitio web y 5. entidades o profesionales de apoyo que contacten a suicidas en redes Sociales.

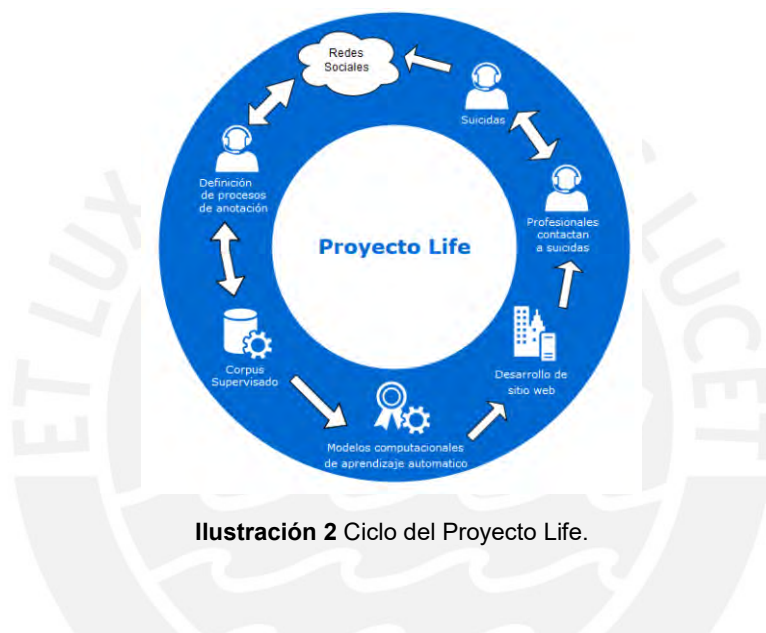


Ilustración 2 Ciclo del Proyecto Life.

Con el presente proyecto doctoral en desarrollo se aportará con las etapas 1, 2 y 3, pero no se podrá determinar el éxito de este proyecto, ya que aún resta una buena etapa de investigación y desarrollo. Teniendo en cuenta, que para que sea exitosa la fase de ejecución de esta plataforma, se necesitara el apoyo de entidades públicas y privadas, con el objeto de que juntos puedan intentar que usuarios suicidas que postean sus intenciones en redes sociales, desistan de las mismas.

1.10. Justificación

El análisis de sentimientos expresados por usuarios de redes sociales constituye una herramienta valiosa porque permite entre otros temas de investigación, detectar problemas de ideación suicida, incitación al suicidio y acoso escolar. Expresados y compartidos con otros usuarios a través de la red en tiempo real, antes y/o mientras están cometiendo el acto de suicidarse.

Actualmente, estas conductas suicidas son estudiadas por grupos de investigación adscritos a universidades, sociedad, gobiernos y entidades de salud internacional. Enfocados en disminuir la tasa de suicidios a nivel mundial, varios de ellos en el marco

del compromiso de sus países con la OMS, para reducirlos en un 10% hasta el año 2020. Suicidios que, en el año 2022, llegaran aproximadamente a 911.375 muertes, y más de 18.227.500 intentos suicidas no consumados [39].

El aprendizaje automático o Machine Learning, es el método supervisado, no supervisado, o semi-supervisado, utilizado por estos grupos de investigadores, para analizar los grandes volúmenes de textos con contenido suicida, extraídos de las redes sociales. Para sin intervención humana y en forma automatizada, aprender a descubrir patrones, tendencias y relaciones entre datos. Pero varias de estas iniciativas no se han ejecutado de la mejor manera porque no existe un Corpus de suicidio estándar, en el cual bazar sus experimentos. Ya que los existentes están basados en notas suicidas, y no son aptos para la detección de mensajes suicidas, sino solo para la diferenciación entre notas suicidas reales y falsas, ya que estos Corpus no tienen ejemplo de mensajes suicidas, ni otro tipo de mensajes que no sea notas suicidas.

El presente estudio pone en evidencia este hecho y establece las directrices, para el desarrollo de un método semi-supervisado que detecte, clasifique y anote en un Corpus de suicidio, de forma automática textos extraídos de redes sociales relacionados con el suicidio. Proceso que permitirá incrementar periódicamente de forma automática el número de textos en sus dos niveles de alerta **Riesgo** y **No Riesgo**. Con los cuales poder entrenar nuevos textos y realizar búsquedas de usuarios suicidas en redes sociales con mayor precisión.

Este trabajo no solo representa una nueva contribución a la ciencia, aportando a las teorías del sector social mediante la formulación y desarrollo de un método semiautomático, para incrementar el número de anotaciones de corpus de ideación suicida a partir de textos de redes sociales, sino que además establece un proceso de evaluación que permitirá de forma automática medir la calidad de los textos previo a su inclusión o exclusión. Esta propuesta forma parte del proyecto Life, cuyo objetivo principal es contribuir a crear una herramienta que automatice la detección de casos de ideación suicida, incitación al suicidio y acoso escolar, a partir de textos escritos por usuarios de redes sociales, en un marco que respete la intimidad de las personas. Para de manera conjunta con organizaciones especializadas, mediante intervenciones oportunas y eficaces basadas en datos científicos, tratamiento y apoyo, incidir en el no suicidio, de miles de personas con ese tipo de ideas.

Por lo expuesto anteriormente, se justifica esta investigación porque desarrollará un método semiautomático para incrementar el número de anotaciones de corpus de ideación suicida a partir de textos de redes sociales. El cual establece un procedimiento estructurado con actividades bien definidas, roles y procesos bien enlazados. Que permitirá a los diversos grupos que investigan, a los usuarios de redes sociales que publican textos relacionados con la ideación suicida, incitación al suicidio y acoso escolar, disponer de un Corpus de suicidio estándar, que sirva como punto de partida para los diversos experimentos que realicen.

1.11. Viabilidad

La viabilidad del presente proyecto de investigación se sustenta en:

1.11.1. Viabilidad técnica

Para la realización del presente proyecto de tesis “**Desarrollo de un método semi-supervisado para detectar, clasificar y anotar en un corpus de suicidio textos extraídos de entornos digitales**”, se cuenta con el apoyo de la compañía iLife (España), Universidad Estatal del Sur de Manabí (Ecuador) y la Pontificia Universidad Católica del Perú.

Los expertos para contactar son:

- Dr. José Manuel Gómez Soriano (**Compañía iLife**)
- Dr. Andrés Melgar Sasieta (**PUCP**)

Además de contar con el apoyo del proyecto “**Tecnologías del lenguaje humano aplicado a la prevención del suicidio en redes sociales**”, de la carrera de Tecnología de la Información, adscrita a la facultad de Ciencias Técnicas de la Universidad Estatal del Sur de Manabí.

Para el desarrollo de esta tesis se van a utilizar herramientas que están a disposición del tesista por parte de:

PUCP. Bases de datos bibliográficas:

- Scopus.
- Web of Science

Software Libre

- Parsifal
- Mendeley
- Weka
- Python
- VOSviemer

1.11.2. Viabilidad económica

La viabilidad económica del proyecto de tesis estará a cargo del programa de becas de la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (Senescyt)

“Programa de Becas para Doctorado (PhD) para Docentes de Universidades y Escuelas Politécnicas” del Ecuador.

1.11.3. Viabilidad temporal

La viabilidad temporal estará enmarcada en el siguiente cronograma de actividades.

PRIMER AÑO			M	A	M	J	J	A	S	O	N	D
Revisión sistemática literaria del estado del arte.			X	X	X	X	X	X	X	X	X	X
Determinación de hipótesis de investigación.					X	X	X					
Determinación de objetivos General y específicos.						X	X					
Alcance, justificación y viabilidad del proyecto.						X	X	X				
Redacción del Plan de Tesis.									X	X	X	
Sustentación pública del Proyecto de investigación doctoral.												X
Inicio de recopilación documental de información para escribir el primer artículo doctoral.			X	X	X	X	X	X	X	X	X	X
SEGUNDO AÑO	E	F	M	A	M	J	J	A	S	O	N	D
Identificación de técnicas de aprendizaje automático.			X	X	X							
Elaboración de marco conceptual, estado del arte.					X	X	X	X	X	X	X	X
Realización de experimentos.							X	X	X	X	X	X
TERCER AÑO	E	F	M	A	M	J	J	A	S	O	N	D
Realización de experimentos.			X	X	X	X	X	X	X	X	X	X
Redacción de primer artículo científico en base a los experimentos realizados.			X	X	X	X	X	X	X	X		
Redacción del documento de tesis.			X	X	X	X	X	X	X	X	X	X
Publicación de primer artículo científico en base a los experimentos realizados.											X	X
CUARTO AÑO	E	F	M	A	M	J	J	A	S	O	N	D
Realización de experimentos.			X	X	X	X	X	X	X	X	X	X

Redacción de segundo artículo científico en base a los experimentos realizados.			X	X	X	X	X	X	X	X		
Finalizar la redacción del documento de tesis.			X	X	X	X	X	X	X	X	X	X
Publicación del segundo artículo científico en base a los experimentos realizados											X	X



Capítulo II

2. Marco conceptual

Los conceptos relacionados que se deben tener en cuenta para el desarrollo del presente proyecto de tesis son los siguientes:

2.1. Suicidio

De acuerdo al diccionario de la real academia española **es la acción y efecto de suicidarse** [136]. Es el acto de quitarse deliberadamente la propia vida al que llegan las víctimas de suicidio, luego de pasar por un proceso de profundo sufrimiento personal con frecuencia en silencio [137].

Entre las personas que corren un mayor riesgo de suicidarse tenemos: quienes anteriormente ya han intentado suicidarse; quienes se sienten desesperanzados, inútiles, agitados, aislados de la sociedad; aquellos que están pasando por situaciones estresantes como la pérdida de un ser querido, servicio militar, separación, problemas financieros o legales; problemas de consumo de alcohol, drogas; pensamientos suicidas y acceso a armas de fuego; trastornos psiquiátricos no diagnosticados, como depresión grave, trastornos postraumáticos o trastorno bipolar; antecedentes familiares de trastornos mentales, abuso de sustancias, suicidio o violencia; quienes tienen enfermedades que se pueden asociar con la depresión y los pensamientos suicidas, como una enfermedad crónica o una enfermedad terminal; las personas lesbianas, gay, bisexuales o transgénero que no cuentan con apoyo familiar o están expuestos a un entorno hostil [29][39].

2.2. Prevención de suicidio.

De acuerdo al diccionario de la real academia española la prevención **es la acción y efecto de prevenir** que algo ocurra [138], en este caso un suicidio. Cada año más de 883.500 personas se suicidan y más de 17.670.000 lo intentan [38]. Por ello, la OMS reconoce al suicidio como una prioridad mundial de salud pública. Y sus estados miembros se han comprometidos a disminuir la tasa de suicidios en un 10% antes del año 2020, implementando estrategias de prevención de suicidios, sistemas de información eficientes, investigaciones universitarias colaborativas sobre salud mental, entre otras [39].

Entre las estrategias de prevención se tienen tres:

De **prevención “universal”**, diseñadas para llegar a toda una población por medio de incrementar la atención de salud, promover la salud mental, reducir el consumo nocivo de alcohol, limitar el acceso a los medios utilizables para suicidarse o promover una información responsable por parte de los medios de difusión [38].

De **prevención “selectiva”**, dirigida a grupos vulnerables como los compuestos por personas que han padecido traumas o abuso, los afectados por conflictos o desastres, los refugiados y migrantes, familiares de suicidas, mediante servicios de ayuda especializados [38].

De **prevención “indicadas”**, las cuales se dirigen a personas vulnerables específicas, mediante el apoyo de la comunidad, el seguimiento a quienes salen de los establecimientos de salud, la capacitación del personal de salud y una mejor identificación y manejo de los trastornos mentales y por uso de sustancias [38].

Estas estrategias son desarrolladas por el estado de manera conjunta o individual con organizaciones especializadas. Las que cuando detectan a tiempo a posibles suicidas, mediante intervenciones oportunas y eficaces basadas en datos científicos, tratamiento y apoyo, intentan incidir para que estas desistan de quitarse la vida [38]. Entre estas organizaciones se encuentran, las organizaciones no gubernamentales de voluntariado sin ánimo de lucro Teléfono de la Esperanza [139], The Samaritans [140] y Asociación la Barandilla [141].

2.3. Corpus.

De acuerdo al diccionario de la real academia española **es el conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etc., que pueden servir de base a una investigación** [142]. Entonces un Corpus lingüístico es un conjunto de textos hablados u escritos debidamente recopilados y posteriormente anotados [71] de forma supervisada (Datos etiquetados manualmente), No supervisada (Datos no etiquetados manualmente), o una combinación de ambos (semi-supervisado) [72][74].

Son utilizados en procesos de investigación, en prácticamente cualquier campo del conocimiento, en que se dese trabajar con una cantidad más o menos grande de muestras de elementos que formen parte de la realidad que se quiere observar. Facilita las tareas mecánicas de recopilación y organización de textos, aportando resultados que solo es posible obtener por medio del procesamiento del lenguaje natural (PLN), tanto para el análisis lingüístico como para el aprendizaje automático [143][144].

Una de las principales contribuciones de la lingüística de Corpus se encuentra en el área de exploración de patrones de uso del lenguaje [145][146]. Se pueden clasificar como: monolingüe, que representa un idioma; comparable, utilizando múltiples Corpus monolingües para crear un marco comparativo; Corpus paralelos, donde se considera el Corpus de un idioma y los datos obtenidos se traducen a otros idiomas [127], [147].

Entre los Corpus existentes más conocidos se tiene: al British Nacional Corpus (BNC) [148], Corpus of Contemporary American English (COCA) [149], Brown Corpus [150], y el Helsinki Corpus of English Texts [4].

2.3.1. Corpus supervisado.

Un Corpus supervisado, está formado por textos reales de un campo específico del conocimiento, anotados de forma manual. Su calidad es medida mediante metodologías estadísticas, que por lo general priorizan los acuerdos o coincidencias entre anotadores [75]. El estadístico Cohen's Kappa, es la metodología estadística más utilizada por los diferentes grupos de investigación, para evaluar los acuerdos entre anotadores en el proceso de anotación de textos a sus Corpus supervisados [151][76][77][78].

En general es una medida sólida, que toma en cuenta tanto el cálculo del porcentaje del acuerdo, como el acuerdo que se produce por casualidad. Conduce a clasificaciones más precisas y ha sido ampliamente utilizada principalmente en la construcción de Corpus supervisados, con textos provenientes de redes sociales, ciencias, biología y ciencias médicas [75]. Otros grupos de investigación en el ámbito de sistemas expertos, comunidades de aprendizaje automático y minería de datos, etc., para medir sus acuerdos en la anotación de texto entre sus evaluadores han utilizado otros métodos como: Yule's Y (1912), Bennett et al's S (1954), Scott's π (1955), Osgood's index (1959), Cohen's κ (1960), Holsti's coefficient (1969), Maxwell's RE (1977), Krippendorff's α (1980), Perreault & Leigh's Ir (1989), Feinstein & Cicchetti's ψ (1990), Potter & Levine-Donnerstein's redefined π (1999), and Gwet's γ (2006) [152].

En el proceso de construcción de un Corpus supervisado o Gold Standard, se involucra una gran cantidad de personas, software y hardware, por esto son costosos computacionalmente y es una tarea que requiere mucho tiempo para su desarrollo. Constituye un verdadero cuello de botella para el desarrollo de muchos enfoques basados en datos para el Procesamiento del Lenguaje Natural (PLN), siendo una de sus principales funciones el entrenamiento de nuevos datos [153].

Entre los Corpus supervisados que han sido creados se tiene: El Corpus IxaMed-GS, compuesto por historias clínicas reales redactadas en español y anotadas manualmente por expertos en farmacología y farmacovigilancia, cuya evaluación de concordancia entre anotadores llegó a 90.53% [154].

2.3.2. Corpus semi-supervisado.

Un Corpus semi-supervisado por lo general está compuesto de una cantidad fija de datos anotados de forma supervisada y datos sin etiquetar anotados de forma semi-supervisada al Corpus, previo a determinar su calidad de forma automática. Con el objeto de incrementar su tamaño, para aumentar el rendimiento del modelo, aunque su calidad pueda disminuir [155]. El proceso de anotación automática de nuevos textos a un Corpus se basa en el proceso de anotación supervisada, pero esta vez de forma automática. Esto permite que el costo computacional por nuevas anotaciones realizadas sea menor, al reducir la carga de trabajo del equipo de anotación, en cuanto a tiempo y recursos invertidos [155][156].

Entre los Corpus que han recibido anotaciones semi-supervisadas se tiene: El Corpus desarrollado por O'Dea y su equipo de trabajo, quienes investigaron si el nivel de preocupación por publicaciones relacionadas con el suicidio en Twitter podría determinarse basándose únicamente en el contenido de la publicación [25].

2.4. Corpus paralelos

Actualmente, la falta de datos anotados en Corpus sobre suicidio en idiomas diferentes al inglés representa un obstáculo para el desarrollo de aplicaciones de PLN, en esta área. Por ello los esfuerzos para reducir la necesidad de anotaciones de roles semánticos se ha centrado principalmente en idiomas distintos al inglés. La proyección de anotaciones es un marco popular para transferir anotaciones semánticas de un idioma a otro, mediante equivalencias estructurales y de traducción presentes en los Corpus paralelos [156].

Los Corpus paralelos son una realidad a partir de la traducción automática y la necesidad que tienen muchos grupos de investigación alrededor del mundo de realizar procesos de investigación, en el campo del procesamiento del lenguaje natural en diferentes idiomas [126][156]. Mediante recursos lingüísticos y anotaciones de alta calidad en la explotación de Corpus paralelos alineados, lo que supone que, si un texto en un idioma ha sido anotado y su traducción no, las anotaciones pueden transferirse, del texto fuente, al destino utilizando la alineación de palabras como puente. Permitiendo la explotación de Corpus existentes principalmente en idioma inglés, para iniciar la creación de Corpus anotados en nuevos idiomas (con pocos recursos) con un esfuerzo humano muy reducido [73][145].

Entre los Corpus que han sido objeto de traducción paralela se tiene: El Corpus FrameNet [144], del Instituto Internacional de Ciencias de la Computación de la universidad de Berkeley, el cual ha sido objeto de diversas traducciones paralelas a diferentes idiomas [156].

2.5 Análisis de sentimientos

Una característica importante del ser humano es conocer los sentimientos y las opiniones de las demás personas a su alrededor acerca de un tema específico, necesidad que, con el desarrollo actual de las comunicaciones y el internet, se ha globalizado [65].

Estos sentimientos y opiniones los encontramos en las redes sociales, blogs, grupo de noticias, etc., en forma escrita, hablada, de imágenes o videos [56][57][58][59], a través de los cuales las personas expresan sus pensamientos, sentimientos, creencias, comportamientos y personalidad [60]. Estos a la vez pueden ser recopilados, procesados y analizados por medio del Procesamiento del Lenguaje Natural, para con la información resultante predecir resultados [61].

Hay que tener en cuenta que el éxito del análisis de sentimientos [63][64][65][66], depende en gran medida de la calidad de los datos implícitos [67] y no específicos [68] que se pueden extraer del gran flujo de información ahora disponible [69][63][70], en las redes sociales [63]. Con el objeto de buscar identificar los puntos de vista subyacentes a un espacio de texto, por medio de metodologías de aprendizaje automático como, por ejemplo, la categorización de texto solo a las porciones subjetivas del documento [66].

Capítulo III

3. Estado de la Cuestión.

En este capítulo se efectuará una revisión sistemática literaria respecto al tema de investigación propuesto, **“MÉTODO SEMI-SUPERVISADO PARA DETECTAR, CLASIFICAR Y ANOTAR EN UN CORPUS DE SUICIDIO TEXTOS EXTRAIDOS DE ENTORNOS DIGITALES”**. La revisión sistemática es una metodología específica en el marco de la ingeniería del software [157]. Que ubica los estudios existentes, selecciona y evalúa contribuciones, analiza y sintetiza datos, e informa la evidencia de tal manera que permite llegar a conclusiones razonablemente claras sobre lo que se conoce y lo que no se conoce de un tema específico, a partir de una pregunta claramente especificada. Se diferencia de otros métodos de revisión debido a sus principios distintos y rigurosos [158] [159][160].

3.1. Definición del Proceso de Revisión Sistemática.

Con esta revisión sistemática, se pretende conocer y evaluar de forma analítica la contribución de estudios existentes si los hubiere. Además de determinar vacíos de conocimiento en el marco de sentimientos suicidas expresados en redes sociales, mediante la selección, evaluación y compilación de resultados de estudios primarios, expresados en forma de artículos científicos, con el fin de proponer un curso de investigación en alguno de los vacíos de conocimientos encontrados [158].

El desarrollo de esta revisión sistemática tuvo como herramienta de soporte, la plataforma online Parsifal, la cual utiliza una metodología tal, que para su construcción necesita se definan los siguientes elementos:

Título de la revisión.

Descripción.

Objetivos.

Definición de criterios PICOC

Preguntas de investigación.

Palabras claves y sinónimos.

Cadena de búsqueda.

Fuentes de información

Criterios de inclusión y exclusión.

Lista de evaluación de la calidad.

Formulario de extracción de datos.

Para seguidamente proceder a cargar en la plataforma, los metadatos de los artículos seleccionados en las bases de datos bibliográficas, que en este caso son:

Documentos generados por la cadena de búsqueda en la base de datos Web of Science.

Documentos generados por la cadena de búsqueda en la base de datos Scopus.

Una vez cargados los artículos científicos de forma manual por parte del investigador, en base a los criterios de inclusión y exclusión, determinar aquellos documentos que han sido:

Aceptados

Rechazados

Duplicados

Para finalmente en base a la información de los artículos científicos aceptados y descrita en los formularios de extracción de datos, proceder a analizar y sintetizar los datos obtenidos, con el fin de obtener resultados y conclusiones razonables.

3.1.1. Herramientas de apoyo para la revisión sistemática.

Entre las herramientas Open Source y de pago, utilizadas en la realización de la revisión sistemática de este proyecto de tesis constan:

3.1.1.1. Bases de datos bibliográficas:

Las bases de datos bibliográficas utilizadas en la presente investigación son:

Scopus⁷. – Base de datos bibliográfica de resúmenes, citas de artículos y revistas científicas producido por Elsevier Co. [161].

Web of Science⁸. – Base de datos bibliográfica de resúmenes, citas de artículos y revistas científicas producido por Clarivate Analytics [162].

⁷ <https://www.scopus.com>

⁸ <https://login.webofknowledge.com>

Google Scholar⁹. – Buscador de Google enfocado en la búsqueda especializada de contenidos bibliográfico científico-académico [163].

3.1.1.2. Gestor bibliográfico.

El gestor bibliográfico utilizado para capturar los metadatos de los artículos científicos incluidos en esta investigación es Mendeley. El cual es una herramienta propietaria y gratuita de captura de metadatos de documentos basada en ambiente web y de escritorio, que integra la gestión de artículos de investigación con funciones para colaborar con investigadores a nivel local y mundial [164]. Las referencias pueden ser utilizadas por el usuario para la construcción de bibliografías en diferentes documentos que este dese generar, en procesadores de texto tanto de ambiente de escritorio como Microsoft Word, o en ambiente online como Overleaf¹⁰.

3.1.1.3. Parsifal.

Parsifal¹¹ es una herramienta Open Source en línea, diseñada para apoyar a los investigadores a realizar revisiones sistemáticas de la literatura en el contexto de la ingeniería del software, los cuales sin importar su ubicación geográfica podrán trabajar juntos en un ambiente de trabajo compartido, diseñando el protocolo de investigación y llevando a cabo la investigación de forma totalmente documentada [160][165].

La metodología de Parsifal cubre los siguientes aspectos: objetivos, términos PICOC, preguntas de investigación, cadena de búsqueda, palabras claves y sinónimos, selección de fuentes documentales y criterios de inclusión y exclusión, listas de verificación y formularios de extracción de datos. Teniendo como limitante reportes de la información procesadas muy restringidos [166].

3.2. Desarrollo de la revisión sistemática literaria Prevención de Suicidios en Redes Sociales.

El tema de la presente revisión sistemática literaria es “**Prevención de suicidios en redes sociales**”.

3.2.1. Objetivos.

Los objetivos de la revisión sistemática son:

⁹ <https://scholar.google.com>

¹⁰ <https://www.overleaf.com>

¹¹ <https://www.parsif.al>

1. Identificar en las bases de datos bibliográficas Scopus y Web of Science por medio de una cadena de búsqueda de palabras claves, estudios primarios referentes a la prevención o ideación de suicidios en redes sociales.
2. Realizar en base a los criterios de inclusión o exclusión determinados en la metodología, el proceso de incluir o excluir de la investigación artículos científicos identificados en las bases de datos bibliográficas.
3. Analizar y sintetizar información relevante de los artículos incluidos en la presente investigación.

3.2.2. Definición de preguntas de investigación.

La definición de las preguntas de investigación fue un punto crítico de la presente metodología de revisión sistemática literaria. Ya que en base a ellas definimos el conocimiento que queremos alcanzar, previo a iniciar el proceso de búsqueda y extracción de información desde estudios primarios que respondieran a las interrogantes planteadas. Para de esta forma poder determinar sus tendencias, grupos de investigación activos, trabajos futuros, etc.

En este marco se planteó las siguientes preguntas de investigación:

PI1: ¿Quiénes son los principales investigadores en el campo de la prevención de suicidios en redes sociales?

PI2: ¿Qué métodos se han utilizado en los procesos de investigación en el campo de la prevención de suicidios en redes sociales?

PI3: ¿Cuáles han sido los conjuntos de datos que han utilizado los grupos de investigación para la realización de sus experimentos en el campo de la prevención de suicidios en redes sociales?

PI4: ¿Qué medidas de evaluación estadísticas se utilizaron en el procesamiento de la información?

PI5: ¿Qué dimensiones se utiliza para entrenar el modelo?

PI6: ¿Cuáles son los algoritmos de clasificación más utilizados para el entrenamiento de Corpus?

PI7: ¿Cuáles son los vacíos de conocimientos existentes en el campo de la prevención de suicidios en redes sociales?

En base a la metodología planteada y una vez definidas las preguntas de investigación en la búsqueda de conceptos específicos se definieron conceptos generales basados en los criterios PICOC (Population, Intervention, Comparison, Outcomes y Context) [157], lo cual permitió identificar estructuras estandarizadas para la formulación de preguntas en el marco de prácticas basadas en evidencias [167][168][169], las cuales se presentan en la Tabla 7.

PICOC	
Criterio	Descripción
Población	Usuarios de redes sociales
Intervención	Comportamiento suicida, suicidio prevención, ideación suicida, suicidio
Comparación	Procesamiento del lenguaje natural
Salida	Método de procesamiento
Contexto	Usuarios de redes sociales

Tabla 7. Metodología PICOC.

3.2.3. Palabras claves

Dados los criterios PICOC basados en las preguntas de investigación se determinaron las siguientes palabras claves en idioma inglés, mismas que fueron utilizadas en la cadena búsqueda de información en las bases de datos bibliográficas:

Computer science

Suicidal ideation

Supervised classifiers

Machine learning

Social networks

Automatic classification

Suicide

3.2.4. Cadena de Búsqueda

Para dar respuestas desde estudios primarios a las interrogantes planteadas por las preguntas de investigación, se desarrolló la siguiente cadena estándar de búsqueda bibliográfica de artículos académicos:

("Social Networks" OR "Natural Language Processing") AND ("ideation suicidal" OR "suicidal" OR "suicidal behavior" OR "Suicide prevention")

La cadena de búsqueda basada en palabras claves del tema investigado [170], sufrió ciertas modificaciones de acuerdo a la sintaxis de las bases de datos bibliográficas en que se ejecutó y que permitió encontrar los resultados que a continuación se detallan:

3.2.4.1. Base de datos bibliográfica Web of Science

Al aplicar la cadena de búsqueda en la **plataforma Web of Science**, esta identifico 134 documentos de estudios primarios. Tabla 8.

Cadena de búsqueda	Numero de artículos
(TS=("Natural Language Processing" OR "social networks") AND ALL=("suicidal behavior" OR "Suicide prevention" OR "ideation suicidal" OR suicidal)) AND TIPOS DE DOCUMENTOS: (Article) Índices=SCI-EXPANDED, SSCI, A&HCI, ESCI Período de tiempo=Todos los años	134

Tabla 8. Cadena de búsqueda Web of Science.

De los 134 documentos identificados por la cadena de búsqueda en la base de datos bibliográfica **web of Science**, estos responden a las siguientes categorías: Psiquiatría (60), Psicología multidisciplinaria (24), Salud publica ambiental ocupacional (15), Psicología clínica (12), Ciencias sociales biomédicas (11), Ciencias multidisciplinarias (9), Servicios de política de salud (7), Desarrollo de psicología (6), Psicología (8) y Neurología clínica (5), tal como se muestra en la Ilustración 3.

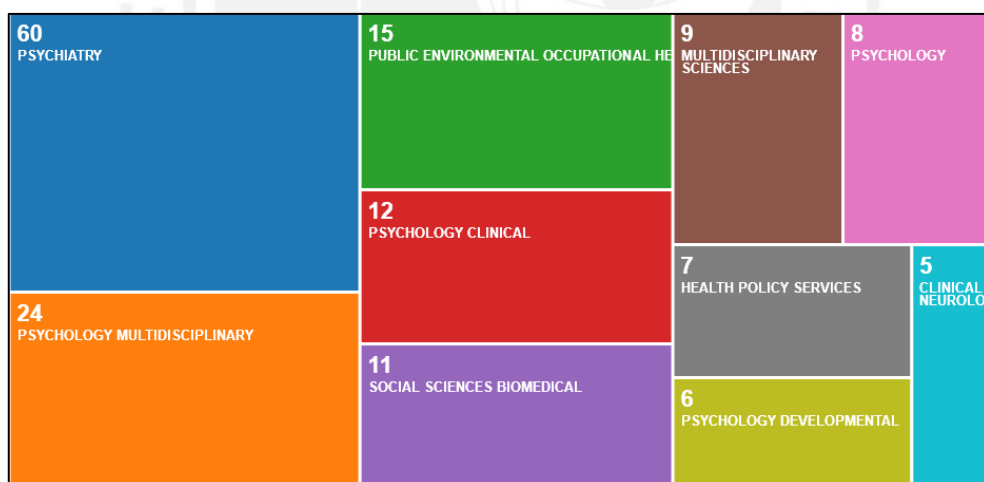


Ilustración 3 Categorías de Web of Science.

Seguidamente con los metadatos de los artículos, se construyó una red bibliométrica de palabras claves con la ayuda de la herramienta VOSviewer¹², la cual permitió conocer entre otros datos los autores con mayor cantidad de citas y publicaciones, tal como se muestra en la tabla 9. Hay que dejar en claro que esta información no necesariamente determina que dichos estudios primarios sean incluidos en el estudio final.

¹² <https://www.vosviewer.com/>

N°	Autor	Documentos	Citaciones
1	Abrutyn, Seth	4	139
2	Mueller Anna S.	4	139
3	Dombrovski, Alexandre Y.	3	108
4	Szanto, Katalin	3	108
5	Baca-Garcia, Enrique	3	68
6	Velupillai, Sumithra	3	28
7	Frey, Laura M.	3	35
8	Hawton, Keith	3	123

Tabla 9. Autores más citados Web of Science.

La herramienta VOSviewer, permitió además construir a partir de los metadatos de los artículos científicos identificados, una red bibliométrica de palabras claves, tal como se muestra en la Ilustración 4. En la cual se pueden observar las palabras claves que se repiten con mayor frecuencia, así como la forma en que están entrelazadas.

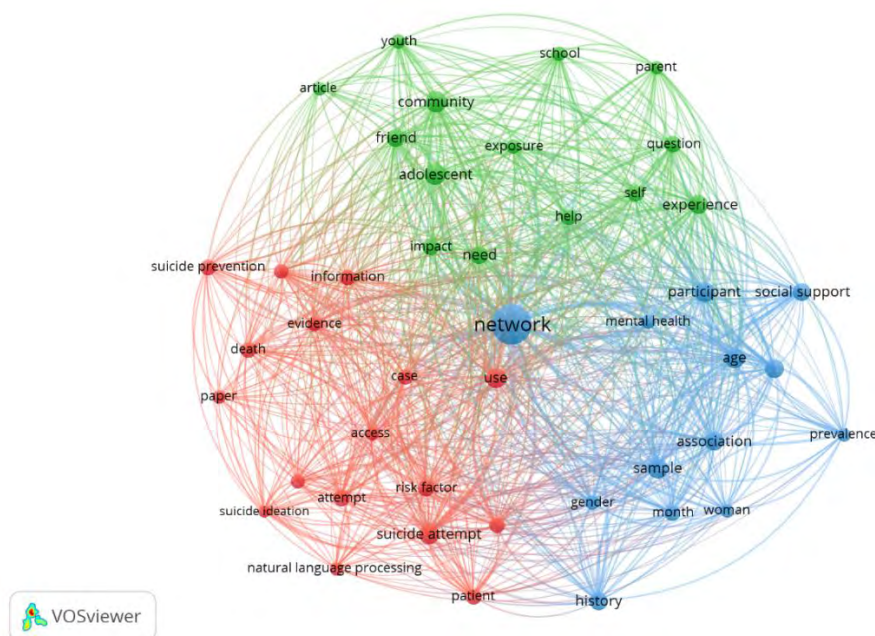


Ilustración 4 Red bibliométrica de palabras claves Web of Science.

3.2.4.2. Base de datos bibliográfica Scopus

De igual manera, al aplicar la cadena de búsqueda en la plataforma **Scopus**, se identificó 224 documentos de estudios primarios. Tabla 10.

Cadena de búsqueda	Numero de artículos
TITLE ("Natural Language Processing" OR "social networks") AND "suicidal behavior" OR "Suicide prevention" OR "ideation suicidal" OR suicidal AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp"))	224

Tabla 10. Cadena de búsqueda Scopus.

De los 224 documentos generados por la cadena de búsqueda en base de datos bibliográfica **Scopus**, estos responden a las siguientes categorías: Medicina (99), Ciencias de la computación (70), Ciencias Sociales (54), Psicología (49), Ingeniería (25), Bioquímica, Genética y Biología Molecular (19), Artes y humanidades (17), Multidisciplinarias (13), Matemáticas (12), Ciencias de la decisión (11), Enfermería (11), Agricultura y Ciencias Biológicas (10), Farmacología, Toxicología y Farmacia (5), Negocios, Gestión y Contabilidad (3), Ciencias Materiales (3), Neurociencias (3), Energía (2), Ciencias Medioambientales (3), Ciencias Materiales (3), Energía (2), Ciencia medioambiental, Física y Astronomía (2), Economía, Econometría y Finanzas (1), Profesiones de la salud (1) e Inmunología y Microbiología (1)), tal como se muestra en la Ilustración 5.

Documents by subject area

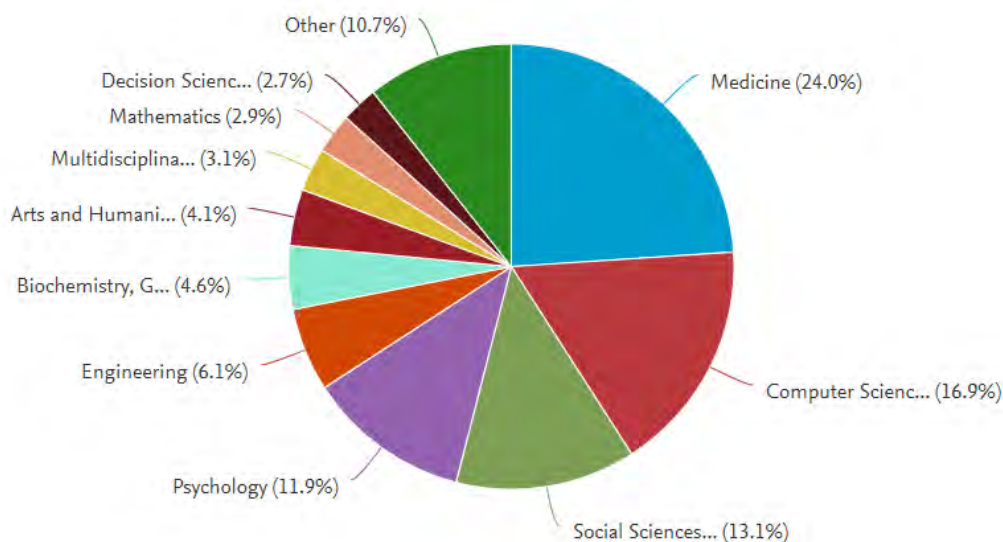


Ilustración 5 Categorías de Scopus.

Seguidamente con los metadatos de los artículos identificados, se construyó la red bibliométrica de palabras claves con la ayuda de la herramienta VOSviewer, lo que permitió tal como sucedió con la información de Web of Science, determinar los autores con mayor número de citas y documentos. Tal como se muestra en la tabla Tabla 11.

N°	Autor	Documentos	Citaciones
1	Fulginiti A.	5	62
2	Rice E.	7	108
3	Tang. J.	8	331

Tabla 11. Autores más citados Scopus.

Además, con la red bibliométrica de palabras se pueden observar las palabras claves que se repiten con mayor frecuencia en los diferentes estudios primarios identificados en esta base de datos, así como la forma en que están entrelazados. Tal como se muestra en la Ilustración 6.

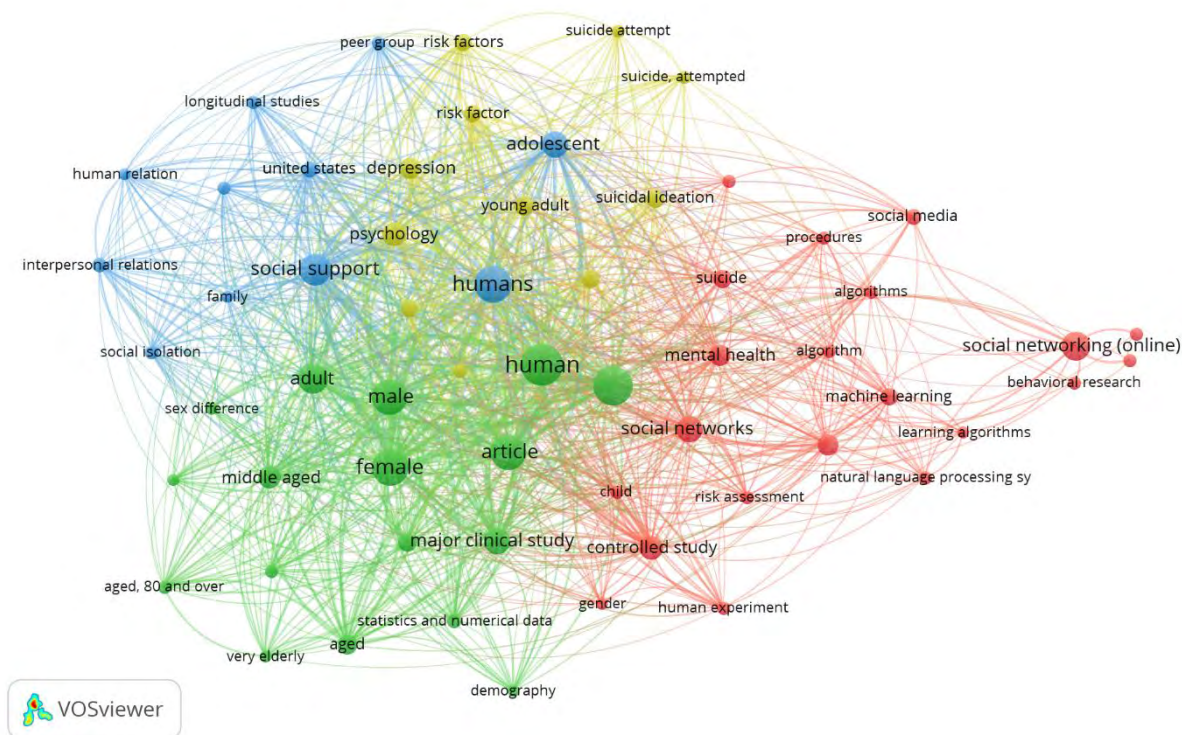


Ilustración 6 Red bibliométrica de palabras claves Scopus.

3.2.5. Criterios de inclusión y exclusión.

Siguiendo con la metodología de revisión sistemática bibliográfica, se determinaron los criterios de inclusión y exclusión que determinaron que estudios se incluyen y cuales se excluyen de la revisión sistemática, tal como se muestra en la Tabla 12. Aunque hay que dejar en claro que a lo largo de la investigación los criterios pueden estar sujetos a cambios a medida que la revisión sistemática avanza a través de las primeras etapas del proceso [171][172].

Criterios de inclusión	Criterio de exclusión
Artículos de ideación suicida	Textos incompletos
Artículos de comportamiento suicida	Artículos de otras áreas del conocimiento
Artículos de prevención de suicidio	Estudios secundarios

Tabla 12. Criterios de inclusión y exclusión.

3.2.6. Lista de verificación de evaluación de la calidad.

Por cada artículo encontrado por la cadena de búsqueda, en las bases bibliográficas e incluido en esta investigación, se evaluó su calidad por medio de la siguiente lista de evaluación de calidad. La cual tiene una valoración por cada ítem de 1 si el resultado de la pregunta es afirmativo y 0 si es negativo, llegando a una calificación máxima de 10 y mínima de 0. La lista de verificación de evaluación de la calidad la podemos ver en la Tabla 13.

N ^a	Preguntas	Se cumple	
		Si	No
1	¿El artículo está relacionado con la prevención o ideación de suicidios en redes sociales?		
2	¿Se describe la hipótesis, finalidad y objeto de estudio?		
3	¿Se describe adecuadamente la metodología?		
4	¿Se describen las características de los individuos incluidos en el estudio?		
5	¿Se justifican la selección de técnicas utilizadas en la investigación?		
6	¿Se describen correctamente los resultados de la investigación?		
7	¿Se miden adecuadamente las variables consideradas por el estudio?		
8	¿Se describe el software utilizado en el presente estudio?		
9	¿Se describen las técnicas estadísticas utilizadas para el análisis de datos?		
10	¿Se evaluó la calidad de los resultados?		

Tabla 13 Listas de verificación de evaluación de la calidad.

3.2.7. Formulario de extracción de datos

El formulario de extracción de datos es un documento valioso en el proceso de esta investigación. Ya que permite extraer información relevante referente a información contenida en los artículos, en base a las preguntas de investigación planteadas. Además de facilitar la obtención de referencias bibliográficas utilizadas por estos autores, lo cual ayuda a profundizar la investigación. Tabla 14.

FORMULARIO DE EXTRACCIÓN DE DATOS			
Nombre del artículo:			
N°	INFORMACION A RECOPILAR	INFORMACION DEL ARTICULO	
1	Nombre de la revista y fecha de publicación.		
2	Investigadores.		
3	Centro de investigación.		
4	¿Hipótesis, finalidad, objetivo de la investigación?		
5	Metodología		
6	¿Qué técnicas de procesamiento del lenguaje natural se utilizaron?		
7	¿Se utilizó el aprendizaje automático supervisado?	Seleccione un campo	No Si
8	¿Qué software de aprendizaje automático y minería de datos se utilizó?	Seleccione un campo	Matlab Weka Python R Otro
9	Se utilizó un Corpus	Seleccione un campo	No Si
10	El Corpus se anotó manualmente	Seleccione un campo	No Si
11	¿Qué algoritmo de clasificación fue utilizado o implementado?		
12	¿Qué características del lenguaje se utilizaron para el algoritmo supervisado?		
13	¿Qué dimensiones se utilizaron para entrenar el modelo?		
14	¿Qué medidas de evaluación se utilizaron?		
15	¿Referencias útiles del artículo?		
16	¿Se han utilizado técnicas semi-supervisado como el bootstrapping?	Seleccione un campo	No Si
17	¿Qué técnicas de semi-supervisado se utilizaron?		
18	¿Qué técnicas se utilizaron para analizar los datos?		
19	¿Cómo se evaluó la calidad de los resultados?		
20	Trabajos futuros		
21	Notas del artículo		

Tabla 14. Formulario de extracción de datos.

3.2.8. Importación de estudios.

Los artículos ubicados a partir de la cadena de búsqueda en la base de datos bibliográfica Scopus fueron 224, y 134 en la base de datos bibliográfica Web of Science. Los metadatos de estos artículos fueron exportados desde estas plataformas por medio de archivos en formatos BitText, los cuales contenían información como: resumen, autor, título, url, año, etc.

3.2.9. Selección de estudios

Una vez cargados en Parsifal los metadatos de los artículos, se procedió a verificar y eliminar los documentos duplicados que en este caso fueron 17. Seguidamente se realizó un análisis de cada uno de los artículos en base a la información contenida en su título, resumen, resultados o en su lectura completa. Para determinar de acuerdo, a los criterios de Inclusión y exclusión, si eran aceptados o rechazados para aplicar o no los formularios de Evaluación de calidad y Extracción de datos. Tabla 15.

Base de datos	Resultados de Búsqueda	Artículos duplicados	Artículos relevantes
Scopus	224	22	7
Web of Science	134		9
Total	368	22	16

Tabla 15. Artículos encontrados.

3.2.9.1. Inclusión de Artículos

En esta sección se detallan los artículos seleccionados en base a los criterios de inclusión y exclusión, que permitieron dar respuestas a las preguntas de investigación planteadas y que forman parte de la bibliografía de este proyecto de tesis y de las publicaciones de tipo Journal que los resultados de este estudio han permitido desarrollar [112][173].

3.2.9.1.1. Artículos incluidos Scopus

Artículos relevantes seleccionados a partir de los documentos encontrados por la cadena de búsqueda desde la base de datos bibliográfica Scopus son. Tabla 16.

Id	Artículo
1	M. Birjali, A. Beni-Hssane, and M. Erritali, "Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Sentiment Prediction in Social Networks," <i>Procedia Comput. Sci.</i> , vol. 113, pp. 65–72, 2017, doi: 10.1016/j.procs.2017.08.290.

2	K. Fu, Q. Cheng, P. W. C. Wong, and P. S. F. Yip, "Responses to a Self-Presented Suicide Attempt in Social Media," http://dx.doi.org/10.1027/0227-5910/a000221 , Jan. 2013, doi: 10.1027/0227-5910/A000221.
3	F. Tapia, C. Aguinaga, and R. Lujé, "Detection of Behavior Patterns through Social Networks like Twitter, using Data Mining techniques as a method to detect Cyberbullying," in <i>2018 7th International Conference On Software Process Improvement (CIMPS)</i> , Oct. 2018, pp. 111–118, doi: 10.1109/CIMPS.2018.8625625.
4	J. Pestian and J. Grupp-Phelan, "A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department," <i>Suicide and life-</i> , 2016, Accessed: Aug. 28, 2017. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/sltb.12180/full .
5	Z. Ivan Santini, A. Koyanagi, S. Tyrovolas, and J. M. Haro, "The association of relationship quality and social networks with depression, anxiety, and suicidal ideation among older married adults: Findings from a cross-sectional analysis of the Irish Longitudinal Study on Ageing (TILDA)," <i>J. Affect. Disord.</i> , vol. 179, pp. 134–141, Jul. 2015, doi: 10.1016/j.jad.2015.03.015.
6	V. M. B. Silenzio, P. R. Duberstein, W. Tang, N. Lu, X. Tu, and C. M. Homan, "Connecting the invisible dots: Reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social networks," <i>Soc. Sci. Med.</i> , vol. 69, no. 3, pp. 469–474, Aug. 2009, doi: 10.1016/j.socscimed.2009.05.029.
7	J. M. Gómez, "Language Technologies for Suicide Prevention in Social Media," <i>Proc. Work. Nat. Lang. Process. 5th Inf. Syst. Res. Work. Days</i> , pp. 21–29, 2014, doi: 10.3115/v1/W14-6903.

Tabla 16. Artículos relevantes Scopus.

3.2.9.1.2. Artículos incluidos Web of Science.

Artículos relevantes seleccionados a partir de los documentos encontrados por la cadena de búsqueda desde la base de datos bibliográfica Web of Science son. Tabla 17.

Id	Artículo
1	B. Desmet and V. Hoste, "Emotion detection in suicide notes," <i>Expert Syst. Appl.</i> , vol. 40, no. 16, pp. 6351–6358, Nov. 2013, doi: 10.1016/J.ESWA.2013.05.050.
2	M. J. Vioules, B. Moulahi, J. Aze, and S. Bringay, "Detection of suicide-related posts in Twitter data streams," <i>IBM J. Res. Dev.</i> , vol. 62, no. 1, Jan. 2018, doi: 10.1147/JRD.2017.2768678.
3	S. Z. Cremades, J. . Gomez Soriano, and B. Navarro-Colorado, "Design, compilation and annotation of a Corpus for the detection of suicide messages in social networks," in <i>Procesamiento de Lenguaje Natural</i> , 2017, vol. 59, pp. 65–72, Accessed: May 28, 2019.
4	V. M. B. Silenzio, P. R. Duberstein, W. Tang, N. Lu, X. Tu, and C. M. Homan, "Connecting the invisible dots: Reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social

	networks,” <i>Soc. Sci. Med.</i> , vol. 69, no. 3, pp. 469–474, Aug. 2009, doi: 10.1016/j.socscimed.2009.05.029.
5	R. W. Acuña Caicedo, J. M. Gómez Soriano, and H. A. Melgar Sasieta, “Assessment of supervised classifiers for the task of detecting messages with suicidal ideation,” <i>Heliyon</i> , vol. 6, no. 8, p. e04412, Aug. 2020, doi: 10.1016/j.heliyon.2020.e04412.
6	G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield, “Analysing the connectivity and communication of suicidal users on twitter,” <i>Comput. Commun.</i> , vol. 73, pp. 291–300, Jan. 2016, doi: 10.1016/J.COMCOM.2015.07.018.
7	N. Masuda, I. Kurahashi, and H. Onari, “Suicide Ideation of Individuals in Online Social Networks,” <i>PLoS One</i> , vol. 8, no. 4, p. e62262, Apr. 2013, doi: 10.1371/journal.pone.0062262.
8	B. O’Dea, S. Wan, P. J. Batterham, A. L. Cleave, C. Paris, and H. Christensen, “Detecting suicidality on Twitter,” <i>Internet Interv.</i> , vol. 2, no. 2, pp. 183–188, May 2015, doi: 10.1016/j.invent.2015.03.005.
9	D. Mowery et al., “Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study.,” <i>J. Med. Internet Res.</i> , vol. 19, no. 2, p. e48, Feb. 2017, doi: 10.2196/jmir.6895.

Tabla 17. Artículos relevantes Web of Science.

3.2.9.2. Artículos relevantes

Una vez concluida la presente revisión sistemática literaria, en base al análisis de cada uno de los artículos, se determinó aquellos que contienen las investigaciones con mayor relevancia en el área de la “**Prevención de suicidios en redes sociales**”. Estos artículos sirvieron de soporte en la presente investigación. Tabla 18.

Id	Artículo
1	J. Pestian and J. Grupp-Phelan, “A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department,” <i>Suicide and life-</i> , 2016, Accessed: Aug. 28, 2017. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/sltb.12180/full .
2	B. Desmet and V. Hoste, “Emotion detection in suicide notes,” <i>Expert Syst. Appl.</i> , vol. 40, no. 16, pp. 6351–6358, Nov. 2013, doi: 10.1016/J.ESWA.2013.05.050.
3	M. J. Vioules, B. Moulahi, J. Aze, and S. Bringay, “Detection of suicide-related posts in Twitter data streams,” <i>IBM J. Res. Dev.</i> , vol. 62, no. 1, Jan. 2018, doi: 10.1147/JRD.2017.2768678.
4	S. Z. Cremades, J. . Gomez Soriano, and B. Navarro-Colorado, “Design, compilation and annotation of a Corpus for the detection of suicide messages in social networks,” in <i>Procesamiento de Lenguaje Natural</i> , 2017, vol. 59, pp. 65–72, Accessed: May 28, 2019.
5	B. O’Dea, S. Wan, P. J. Batterham, A. L. Cleave, C. Paris, and H. Christensen, “Detecting suicidality on Twitter,” <i>Internet Interv.</i> , vol. 2, no. 2, pp. 183–188, May 2015, doi: 10.1016/j.invent.2015.03.005.
6	D. Mowery et al., “Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study.,” <i>J. Med. Internet Res.</i> , vol. 19, no. 2, p. e48, Feb. 2017, doi: 10.2196/jmir.6895.

Tabla 18. Artículos relevantes.

3.2.10. Desarrollo de las preguntas de investigación

Una vez analizada la información recopilada en los formularios de extracción de datos, proveniente de los artículos científicos incluidos en esta investigación, se procedió a responder las preguntas de investigación planteadas en esta revisión sistemática. Información que se presenta a continuación:

3.2.10.1. Primera pregunta

La primera pregunta de investigación plantea la interrogante de conocer, *¿Quiénes son los principales investigadores en el campo de la prevención de suicidios en redes sociales?*

Para dar respuesta a esta pregunta, se presenta en la Tabla 19, información referente a los autores de los artículos incluidos en esta investigación y que de acuerdo con el análisis desarrollado son los más representativos en el área de conocimiento de **“Prevención de suicidios en redes sociales”**.

Artículo	Investigadores	Centro de investigación	País	Año de publicación	Medio de publicación
“Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Prediction in Social Networks”	M. Birjali, A. Beni-Hssane, and M. Erritali	LAROSERI Laboratory, Department of Computer Sciences, University of Chouaib Doukkali, Faculty of Sciences, El Jadida, Morocco	Marruecos	2017	Journal Procedia Computer Science
Responses to a Self-Presented Suicide Attempt in Social Media	K. Fu, Q. Cheng, P. W. C. Wong, and P. S. F. Yip	Centro de Estudios de Periodismo y Medios, Universidad de Hong Kong, Pokfulam, Hong Kong	China	2013	Journal Crisis
Detection of Behavior Patterns through Social Networks like Twitter, using Data Mining techniques as a method to detect Cyberbullying	F. Tapia, C. Aguinaga, and R. Luje	Universidad de las Fuerzas Armadas ESPE, Departamento de Ciencias de la Computación, Quito - Ecuador	Ecuador	2018	7th International Conference On Software Process Improvement (CIMPS)
A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department	J. Pestian and J. Grupp-Phelan	Department of Pediatrics, Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center and Department of	EEUU	2016	Journal Suicide and Life-Threatening Behavior

		Psychiatry and Behavioral Neuroscience, University of Cincinnati, Cincinnati, OH, USA			
The association of relationship quality and social networks with depression, anxiety, and suicidal ideation among older married adults: Findings from a cross-sectional analysis of the Irish Longitudinal Study on Ageing (TILDA)	Z. Ivan Santini, A. Koyanagi, S. Tyrovolas, and J. M. Haro	Parc Sanitari Sant Joan de Déu, Universitat de Barcelona, Fundació Sant Joan de Déu/CIBERSAM	España	2015	Journal of Affective Disorders
Connecting the invisible dots: Reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social networks	Silenzio, Vincent MB and Duberstein, Paul R and Tang, Wan and Lu, Naiji and Tu, Xin and Homan, Christopher M	Center for the Study and Prevention of Suicide, University of Rochester, 300 Crittenden Avenue, Rochester, NY 14642, United States.	EEUU	2009	Journal Social Science & Medicine
Language Technologies for Suicide Prevention in Social Media	J. M. Gómez	Universidad de Alicante Grupo de investigación "Análisis de sentimiento aplicado a la prevención del suicidio en redes sociales (ASAPS)",	España	2014	Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)
Emotion detection in suicide notes	B. Desmet and V. Hoste	Language and Translation Technology Team, University College Ghent, Groot-Brittanniëlaan 45, 9000 Ghent, Belgium. Tel.: +32 9 224 97 53.	Bélgica	2013	Journal Expert Systems with Applications
Detection of suicide-related posts in Twitter data streams	M. J. Vioules, B. Moulahi, J. Aze, and S. Bringay	Laboratorio de Informática, Robótica y Microelectrónica de Montpellier.	Francia	2018	IBM Journal of Research and Development
Design, compilation and annotation of a Corpus for the detection of suicide messages in social networks	S. Z. Cremades, J. . Gomez Soriano, and B. Navarro-Colorado	Universidad de Alicante Grupo de investigación "Análisis de sentimiento aplicado a la prevención del suicidio en redes sociales (ASAPS)",	España	2017	Procesamiento del Lenguaje Natural, 2017, no 59, p. 65-72.
Analysing the connectivity and communication	G. B. Colombo, P. Burnap, A. Hodrog, and J. Scourfield	School of Computer Science and Informatics,	Reino unido	2016	Journal Computer Communications

of suicidal users on twitter		Cardiff University, Queens Buildings, 5 The Parade, Cardiff, United Kingdom			
Suicide Ideation of Individuals in Online Social Networks	N. Masuda, I. Kurahashi, and H. Onari	Department of Mathematical Informatics, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, Japan	Japón	2013	Journal PLOS ONE
Detecting suicidality on Twitter	B. O'Dea, S. Wan, P. J. Batterham, A. L. Caelear, C. Paris, and H. Christensen	Black Dog Institute, The University of New South Wales, Hospital Road, Randwick, NSW 2031, Australia	Australia	2015	Journal Internet Interventions
Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study	D. Mowery	Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States	EEUU	2017	Journal JMIR

Tabla 19. Información de los investigadores.

A su vez de este grupo, sobresalen como autores de los artículos más destacados los autores: José Manuel Gómez Soriano, Bart Desmet, John P. Pestian, Bridianne O'Dea y M. Johnson Vioulès.

3.2.10.2. Segunda pregunta

La segunda pregunta de investigación plantea la interrogante de conocer, **¿Qué métodos se han utilizado en los procesos de investigación en el campo de la prevención de suicidios en redes sociales?**

Para dar respuesta a esta pregunta, luego del análisis de los formularios de extracción de datos, se concluyó: Tabla 20.

Población.	La población sujeta de estudio en los artículos científicos incluidos en la investigación han sido usuarios de redes sociales
Fuentes de información.	La fuente información mayormente utilizada ha sido la red social Twitter, pero además han existido proyectos de investigación que han utilizado fuentes de información como MySpace, Facebook, Reddit, Weibo, Instagram, blogs y fórums.
Herramientas de extracción de información.	Para permitir la extracción de información desde las redes sociales, se han utilizado cuentas de desarrolladores y usuarios, estas cuentan por lo general con una API, la cual es un interfaz de programación entre la red social y las aplicaciones de descarga.
Herramientas de procesamiento de información.	Para el procesamiento de la información se han utilizado herramientas como Weka, R, Python, GATE Developer.

Tabla 20. Métodos

3.2.10.3. Tercera pregunta

La tercera pregunta de investigación plantea la interrogante de conocer, *¿Cuáles han sido los conjuntos de datos que han utilizado los grupos de investigación para la realización de sus experimentos en el campo de la prevención de suicidios en redes sociales?*

Los conjuntos de datos utilizados en los diferentes experimentos realizados por los grupos de investigación cuyos artículos han sido incluidos en esta revisión sistemática, provienen de las redes sociales. Mismos que luego de su recopilación, fueron procesados por medio del aprendizaje automático o Machine Learning, pasando por un proceso de etiquetado automático, manual o mixto en la conformación de sus conjuntos de datos o Corpus [96][97][84][98].

Entre estos conjuntos de datos se encuentran el Corpus Life [96], Corpus SAD [87] y otros conjuntos de datos sin nombres específicos.

3.2.10.4. Cuarta pregunta

La cuarta pregunta de investigación plantea la interrogante de conocer, *¿Qué medidas de evaluación estadísticas se utilizaron en el procesamiento de la información?*

Entre las medidas de evaluación estadística utilizadas en los diferentes proyectos de investigación se tiene: Tabla 21.

Medidas de evaluación	Descripción
ANOVA	Técnica de la varianza, también conocida como análisis factorial, se utiliza para el estudio del efecto de uno o más factores (cada uno con dos o más niveles) sobre la medida de una variable continua.
Chi Square	Se utiliza para probar relaciones entre variables categóricas.
Cohen's kappa	Metodología de acuerdos que ajusta el efecto al azar entre evaluadores en la proporción de la concordancia observada para elementos cualitativos obtenidos de redes sociales y foros. Utilizada por todos los grupos de investigación analizados en esta revisión y que entre sus experimentos estuvo la creación de un Corpus supervisado.
Método de Montecarlo	Método no determinista o estadístico numérico, usado para aproximar expresiones matemáticas complejas y costosas de evaluar con exactitud.
Regresión Logística	Tipo de análisis de regresión utilizado para predecir el resultado de una variable categórica en función de las variables independientes o predictoras.
T-test	Prueba en la que el estadístico utilizado tiene una distribución t de Student si la hipótesis nula es cierta.

F-score	Mediada de precisión que combina precision y recall, en donde una puntuación de F1 alcanza su mejor valor en 1 y peor puntuación en 0.
Recall	Permite determinar la cantidad que el modelo de machine Learning es capaz de identificar.
Precision	Mide el número de veces que el sistema ha identificado un tema correctamente, entre todas las veces que ese tema ha sido identificado por el sistema.

Tabla 21. Medidas de evaluación estadísticas.

3.2.10.5. Quinta pregunta

La quinta pregunta de investigación plantea la interrogante de conocer, *¿Qué dimensiones se utiliza para entrenar el modelo?*

Las dimensiones utilizadas en los diferentes proyectos incluidos en esta investigación han sido el punto de partida en la búsqueda de textos en las redes sociales. Entre las dimensiones utilizadas tenemos:

Depresión, ansiedad, suicidio, ideación suicida, burla, denigración, insulto, acoso, intimidación, culpa, miedo, desesperanza, dolor, insomnio, aturcido, matar, muerte, morir, divorcio, abuso, negligencia, despedido, desempleado, vagabundo, agonía, llorar, derrota, deprime, privación, desesperación, devastación, desventaja, decepcionar, desanimarse, descorazonarse, desilusión, insatisfecho, fatalidad, aburrido, fallar, tumba, pena, sombrío, nostalgia, desesperada, herida, inadecuada, inferior, aislado, perder, melancolía, llorar, lástima, triste, sollozo, dolor, sufrimiento, lágrimas, tragedia, infeliz, sin importancia, sin éxito, inútil.

3.2.10.6. Sexta pregunta

La sexta pregunta de investigación plantea la interrogante de conocer, *¿Cuáles son los algoritmos de clasificación más utilizados para el entrenamiento de Corpus?*

Entre los algoritmos de aprendizaje utilizadas en los diferentes proyectos incluidos, y utilizados de manera conjunta con los Corpus desarrollados, tenemos: Tabla 22.

Algoritmo	Descripción
Support Vector Machines (SVMs)	Algoritmo de aprendizaje supervisado, están relacionados con problemas de clasificación y regresión. Su uso es preferido por los diversos grupos de investigación.
Logistic Regression	Algoritmo de clasificación utilizado para predecir la probabilidad de una variable dependiente categórica.
BayesNet	Algoritmo grafo probabilístico que representa un conjunto de variables aleatorias y sus

	dependencias condicionales a través de un grafo acíclico dirigido.
SimpleLogistic	Algoritmo de regresión logística utilizado para predecir la probabilidad de una variable dependiente categórica.
SMO	Algoritmo que forma parte de la plataforma weka.
IBK	Algoritmo que forma parte de la plataforma weka.
Kstar	Algoritmo que determina cuales son las instancias más parecidas, puede utilizar la entropía, o contenido de información de las instancias, como medida de distancia entre ellas. Este algoritmo fue en un grupo de experimentos el que obtuvo mejores resultados.
AdaBostM1	Algoritmo que forma parte de la plataforma weka.
Bagging	Algoritmo que forma parte de la plataforma weka.
CVParamSelection	Algoritmo que forma parte de la plataforma weka.
MultiClassifier	Algoritmo que forma parte de la plataforma weka.
MultiClassUp	Algoritmo que forma parte de la plataforma weka.
MultiSchema	Algoritmo que forma parte de la plataforma weka.
RandomCommittee	Algoritmo que forma parte de la plataforma weka.
RandomFiltClass	Algoritmo que forma parte de la plataforma weka.
RandomSubSpace	Algoritmo que forma parte de la plataforma weka.
Stacking	Algoritmo que forma parte de la plataforma weka.
Vote	Algoritmo que forma parte de la plataforma weka.
WeightedDistances	Algoritmo que forma parte de la plataforma weka.
InputMappedClassifier	Algoritmo que forma parte de la plataforma weka.
DecisionTable	Algoritmo que forma parte de la plataforma weka.
Jrip	Algoritmo que forma parte de la plataforma weka.
OneR	Algoritmo que forma parte de la plataforma weka.
PART	Algoritmo que forma parte de la plataforma weka.
ZeroR	Algoritmo que forma parte de la plataforma weka.
HoeffdingTree	Algoritmo que forma parte de la plataforma weka.
J48	Algoritmo que forma parte de la plataforma weka.
LMT	Algoritmo que forma parte de la plataforma weka.
RandomForest	Algoritmo que forma parte de la plataforma weka.
RandomTreede	Algoritmo que forma parte de la plataforma weka.

Tabla 22. Algoritmos de clasificación

3.2.10.7. Séptima pregunta

La séptima pregunta de investigación plantea la interrogante de conocer, **¿Cuáles son los vacíos de conocimientos existentes en el campo de la prevención de suicidios en redes sociales?**

Entre los vacíos de conocimientos encontrados se encuentran: Tabla 23.

Vacío de conocimiento	Descripción
VC1	No existencia de un sistema o plataforma en línea que se ocupe del proceso de prevención del suicidio en redes sociales.

VC2	No existencia de un método que permita incrementar el tamaño del Corpus de suicidio sin perjudicar su calidad.
VC3	No existencia de una herramienta de validación automática de la calidad de los textos relacionados con el suicidio obtenidos desde entornos digitales.
VC4	No utilización del procesamiento de imágenes para la detección de contenido suicida.

Tabla 23. Vacíos de conocimiento

3.3. Desarrollo de metodología de Revisión Sistemática Literaria Análisis de Sentimientos en Redes Sociales.

Para una mayor comprensión de los sentimientos de los usuarios expresados en redes sociales, se realizó a más de la revisión sistemática **Prevención de suicidios en redes sociales**, la revisión sistemática **Análisis de sentimientos en redes sociales**, que permitió conocer de una forma amplia trabajos científicos realizados en el campo de sentimientos expresados por los usuarios en redes sociales de manera general. En la revisión sistemática de **Análisis de Sentimientos en Redes Sociales** se determinó la siguiente información:

3.3.1. Preguntas estructuradas PICOC

Las preguntas estructuradas PICOC de esta investigación se las puede ver en la Tabla 24.

PICOC	
Población	Usuarios de redes sociales
Intervención	Análisis de sentimientos, redes sociales
Comparación	Procesamiento del lenguaje natural
Salida	Método de procesamiento
Contexto	Usuarios de redes sociales

Tabla 24 Preguntas estructuradas PICOC.

3.3.2. Cadena de Búsqueda.

Para esta revisión sistemática, se desarrolló la cadena estándar de búsqueda bibliográfica de artículos académicos (TITLE-ABS-KEY ("sentiment analysis") AND TITLE-ABS-KEY ("social networks")) AND (LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) AND (LIMIT-TO (ACCESSTYPE(OA))).

3.3.2.1. Base de datos bibliográfica Web of Science.

Al aplicar la cadena de búsqueda en la **plataforma Web of Science**, esta encontró 226 documentos de estudios primarios. Tabla 25.

Cadena de búsqueda	Numero de artículos
TEMA: ("sentiment analysis") AND TEMA: ("social networks") Refinado por: TIPOS DE DOCUMENTOS: (ARTICLE) Período de tiempo: 2018-2020. Índices: SCI-EXPANDED, SSCI, A&HCI, ESCI.	226

Tabla 25 Cadena de búsqueda Web of Science.

Seguidamente con los metadatos de los artículos, se construyó una red bibliométrica de palabras claves con la ayuda de la herramienta VOSviewer, la cual permitió conocer entre otros:

Lista de autores más citados. Tabla 26.

Nº	Autor	Documentos	Citaciones
1	Herrera-Viedma, E.	5	106
2	Morente=Molinera, J. A.	5	106
3	Ramon Saura, José	5	40
4	Reyes-Menéndez. Ana	5	30

Tabla 26 Autores más citados Web of Science.

Red bibliométrica de palabras claves, tal como se muestra en la Ilustración 7.

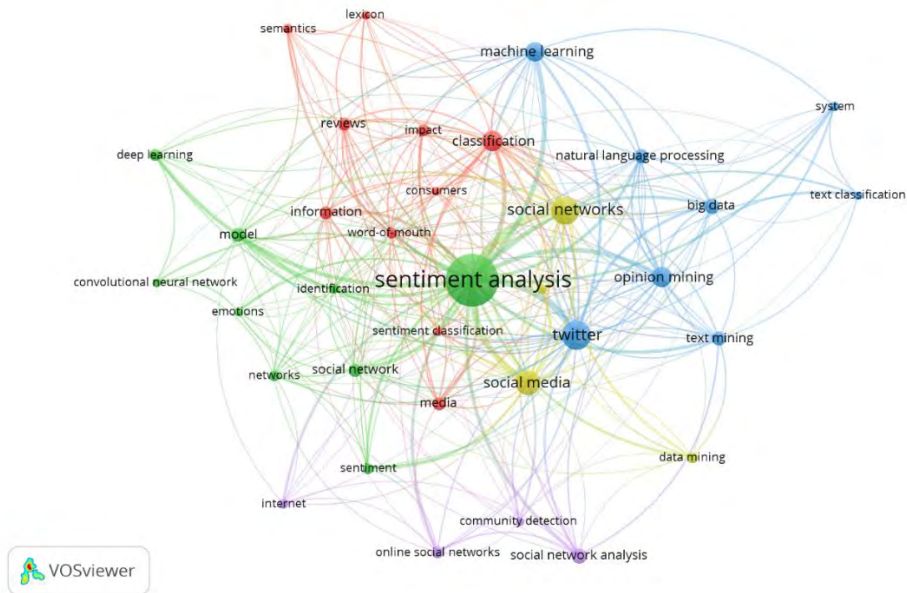


Ilustración 7 Red bibliométrica de palabras claves Scopus.

3.3.2.2. Base de datos bibliográfica Scopus

De igual manera, al aplicar la cadena de búsqueda en la plataforma **Scopus**, se encontró 480 documentos de estudios primarios. Tabla 27.

Cadena de búsqueda	Numero de artículos
TITLE-ABS-KEY ("sentiment analysis") AND ("social networks") AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018)) AND (LIMIT-TO (ACCESSTYPE(OA)))	480

Tabla 27 Cadena de búsqueda Scopus.

Seguidamente con los metadatos de los artículos, se construyó una red bibliométrica de palabras claves con la ayuda de la herramienta VOSviewer, la cual permitió conocer entre otros:

Lista de autores más citados. Tabla 28.

N°	Autor	Documentos	Citaciones
1	Reyes-Menendez, A..	10	10

2	Saura, J. R	13	10
3	Wang, Y	7	0
4	Zhang, Y.	6	0

Tabla 28 Autores más citados Scopus.

Red bibliométrica de palabras claves, tal como se muestra en la Ilustración 8.

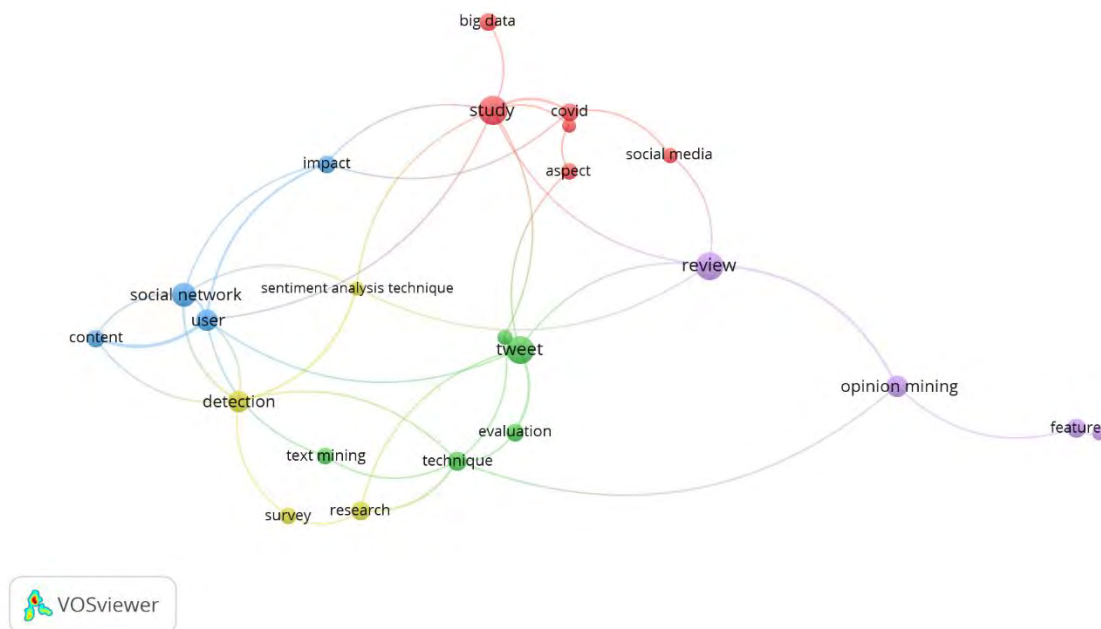


Ilustración 8 Red bibliométrica de palabras claves Scopus.

3.3.3. Criterios de inclusión y exclusión.

Siguiendo con la metodología de revisión sistemática bibliográfica, se seleccionaron los criterios de inclusión y exclusión de los artículos. Tabla 29.

Criterios de inclusión	Criterio de exclusión
Artículos de análisis de sentimientos Artículos de redes sociales	Textos incompletos Artículos de otras áreas del conocimiento Estudios secundarios

Tabla 29 Criterios de inclusión y exclusión.

Una vez cargados en Parsifal los metadatos de los artículos, se eliminaron los documentos duplicados y al igual que en la revisión sistemática anterior se procedió a seleccionar y a descartar artículos basados en los criterios de inclusión y exclusión en base a sus títulos, resumen, resultados o lectura de artículo completos. Seguidamente

se procedió a aplicar el formulario de Evaluación de calidad y Extracción de datos por cada artículo seleccionado, cabe indicar que varios de estos artículos fueron agregados a la bibliografía de esta investigación, con el fin de justificar las afirmaciones realizadas en este documento y en los artículos científicos desarrollados en este proceso doctoral.



Capítulo IV

4. Análisis de métodos para anotación de Corpus

En este capítulo se desarrollará el primer objetivo específico **Proponer un método semi-supervisado que detecte, clasifique y anote en un Corpus de suicidio textos extraídos de entornos digitales**. Para esto analizaremos, documentos académicos relacionados con métodos actuales de anotación automática semi-supervisada, de registros de datos obtenidos desde redes de microblogs, blogs, grupos de noticias, etc., que se encuentran disponibles y han sido ubicados en la **revisión sistemática literaria** realizada en el capítulo III, haciendo énfasis en Corpus de suicidios.

4.1. Evolución de la información

En un mundo en constante evolución, las personas por medio de sus redes sociales se han convertido en consumidores y generadores de información, llegando en promedio a generar 1,7 Mb cada segundo. Aportando con ello de una manera significativa a su crecimiento, que ha pasado de 130 exabytes en el 2005 a 44 zettabytes en el 2020, y llegará a 175 zettabytes en el 2025 [8].

El crecimiento de la información generada en el mundo ha sido posible gracias a la infraestructura de telecomunicación que se ha desarrollado a través de cables de fibra óptica [174], radio, satelital, entre otras [175][176], que ha permitido el auge de las redes sociales [177].

Cabe destacar que de los aproximadamente 44 zettabytes de información que la humanidad ha generado hasta la actualidad, el 33% corresponde a datos etiquetados y el 77% a datos no etiquetados. De los cuales menos del 0,5% está siendo analizado y menos de un 20% está protegido [8]. Generando esto oportunidades para aquellas empresas u usuarios que sean capaces de analizar y extraer valor de la expansión global de los datos [8][9], a través del análisis de los sentimientos expresados en redes sociales [56][57][58][59], bases de datos, entre otros, por medio del aprendizaje de máquina o machine Learning, sea este aprendizaje supervisado [10], no supervisado o semi-supervisado [11][12], aprendizajes posibles de realizar con la ayuda de Corpus Lingüísticos [74].

4.2. Corpus Lingüísticos

Entre las fortalezas de los enfoques basados en Corpus lingüísticos, se tiene que estos proporcionan grandes bases de datos de discursos naturales, para la realización de análisis empíricos de patrones reales de uso en un idioma. Que, al ser combinado con herramientas computacionales, permiten la realización de análisis de un alcance que de otro modo no sería factible, por los grandes volúmenes de datos que se deben procesar [120]. Por ello en casi todos los campos del conocimiento, grupos de investigación han desarrollado Corpus, entre los cuales se tiene a:

O'Connor y su equipo de trabajo, centraron su investigación en la creación de un Corpus estándar de oro supervisado especializado en el abuso de medicamentos, con datos recogidos en idioma inglés de la red social Twitter, mismo que tuvo un acuerdo entre anotadores de 0,86 (Cohen's kappa). Para demostrar la utilidad del Corpus en los sistemas de entrenamiento de clasificación automática, se utilizaron grupos de tweets para entrenar con los algoritmos, Naive Bayes (NB), Random Forest (RF), redes neuronales convolucionales profundas (dCNN) y Support Vector Machines (SVM) que con una precisión de 0,73% obtuvo el mejor desempeño [102].

Sarker y su equipo, trabajaron en la exploración de enfoques de procesamiento del lenguaje natural (NLP), para generar características útiles a partir de textos y utilizarlos juntamente con algoritmos de aprendizaje automático optimizados para la clasificación automática de segmentos de textos asertivos ADR, para la detección de reacciones adversas al abuso de medicamentos. Los datos para los experimentos se recopilaron en la red social Twitter y otras fuentes. El acuerdo entre los anotadores fue de 0,86 (Cohen's kappa). El Corpus generado fue utilizado para entrenar los algoritmos Naive Bayes (NB), Máxima Entropía (ME) y Support Vector Machines (SVM) que con una precisión de 0,81% obtuvo el mejor desempeño [178].

O'Dea y su equipo, trabajaron en detectar si el nivel de preocupación por las publicaciones relacionadas con el suicidio en Twitter podría servir para generar un Corpus de capacitación para modelos de aprendizaje automático e implementar un clasificador informático automatizado que pudiera replicar la precisión de los codificadores humanos. Los datos para los experimentos se obtuvieron de la red social Twitter. La tasa de acuerdo global entre los codificadores humanos fue del 0,76 (Cohen's kappa). Los algoritmos de clasificación utilizados fueron Regresión Logística y Support Vector Machines (SVM) con TF / IDF sin filtro el cual con una precisión de 0,67% obtuvo el mejor desempeño [25].

Gómez y su equipo, trabajaron en la creación del Corpus Life, el cual está orientado a la detección de ideación suicida y está compuesto de textos en idiomas inglés y español. Este Corpus se construyó recuperando textos de varias redes sociales, siendo la tasa de acuerdo global entre los codificadores humanos de 0,52% (Cohen's kappa) [108], con cuatro características.

Cabe señalar que tanto **O'Connor** [102], [108], **Sarker** [178], **O'Dea** [25], y **Gómez** [108], desarrollaron Corpus con una tecnología correcta y un acuerdo de calidad [25][76][77][25], [78][80][81][151][154], medido por del estadístico Cohen's kappa [121], que viene dada por la ecuación 1.

$$k = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)}$$

Ecuación 1. Cohen's kappa

El proceso de desarrollar un método semi-supervisado que detecte, clasifique y anote en un Corpus de forma automática textos extraídos de entornos digitales relacionados con el suicidio, se realizó de acuerdo con el método propuesto en la ilustración 9.

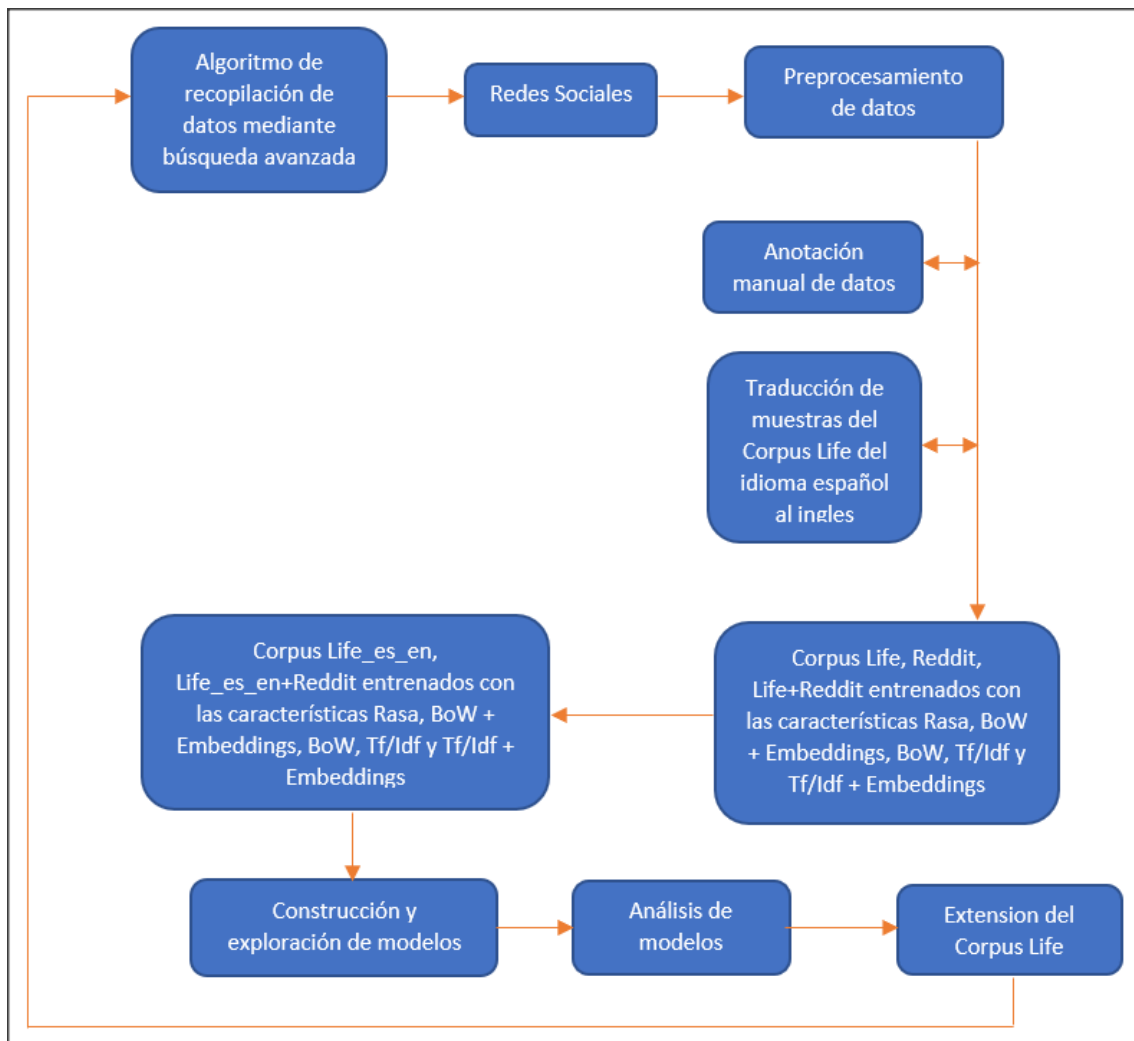


Ilustración 9. Método semi-supervisado.

El desarrollo de estos experimentos consta en el subcapítulo 7.2 **Experimentos 2**, de este documento.

Capítulo V

5. Métodos de evaluación de calidad de anotaciones.

En este capítulo se desarrollará el **segundo objetivo específico** “**Realizar la validación de la calidad de los textos relacionados con el suicidio obtenidos desde entornos digitales**”. Para esto analizaremos los métodos actuales de evaluación de calidad de anotaciones.

5.1. Introducción

Previo a ser incluidos en un Corpus, la calidad de la información contenida en los textos extraídos de las redes sociales debe ser medida, lo que determinara su inclusión o exclusión dada su calidad y validez para la tarea requerida. Evidenciando de esta manera la importante no solo de tener acceso a la información, sino además de saber qué hacer con ella, es decir, saber cómo procesarla adecuadamente y como utilizarla de manera práctica, con el fin de resolver problemas específicos. Lo que ha llevado a los investigadores a replantearse la forma tradicional de realizar el procesamiento de datos, con el fin de que verdaderamente sirvan de base para sustentar sus investigaciones con un enfoque científico [179].

El proceso de creación de Corpus lingüísticos es necesario para poder analizar grandes cantidades de datos, en los cuales la **fiabilidad entre codificadores o anotadores de datos**, es el indicador cualitativo de mayor utilidad para medir la calidad de la información. Este método de evaluación de calidad, es utilizado por diversos grupos de investigación en áreas tan diversas como psicología, sociología, educación, medicina, marketing, computación, etc., para determinar qué información les será útil y cual no [152].

El estudio de la validación de la información se remonta al año 1948, con la publicación del artículo ***The structure and function of communication in society***, por parte de Harold D. Lasswell [180], en el cual se da un gran impulso a la estructuración de los estudios de contenidos de datos, y en el mismo se hace hincapié en que un acto utilizado para medir la calidad de los datos debe responder a las siguientes preguntas:

¿Quién? (Investiga los factores que inician y orientan el acto de la comunicación).

¿Dice que? (Análisis de contenido).

¿Como? (En qué canal)

¿A quién?

¿Con qué efecto?

En épocas más recientes **Neuendorf**, en su libro **Análisis de Contenidos**, hace referencia que la necesidad de profundizar la investigación de contenidos, ha aumentado drásticamente con el inagotable flujo de información disponible en la World Wide Web. Hoy en día la necesidad de verificar la calidad de la información, ha permitido que el **Análisis de Contenidos**, al cual se puede definir como el análisis sistemático, objetivo y cuantitativo de las características de mensajes se haya convertido en una de las técnicas más populares y de rápida expansión para la investigación cuantitativa, apoyada por aplicaciones informáticas que han permitido que el estudio organizado de los mensajes sea más rápido y sencillo [181].

5.2. Pasos del análisis de contenido

Previo a iniciar el análisis de contenidos de los datos cuantitativos obtenidos en el marco de la investigación desarrollada, es imperativo identificar y desarrollar los siguientes pasos [182]:

1. Establecer hipótesis
2. Identificar datos apropiados (texto u otro material comunicativo)
3. Determinar el método de muestreo y la unidad de muestreo
4. Sacar muestra
5. Establecer una unidad de recopilación de datos y una unidad de análisis
6. Establecer un esquema de codificación que permita probar hipótesis
7. Codificar datos
8. Verificar la confiabilidad de la codificación y ajustar el proceso de codificación si es necesario
9. Analizar datos codificados, aplicando pruebas estadísticas apropiadas
10. Escribir resultados

5.3. Índices estadísticos

Para la realización de un **Análisis de contenidos**, existen diversos índices estadísticos en los cuales los investigadores se apoyan para tal fin. En este marco Zhao y su grupo de trabajo [152], analizaron diferentes índices estadísticos, para la realización del Análisis de contenidos de información recopilada en el marco de una investigación, dividiéndolos por sus características en dos grupos, el primero llamado **Índices no Ajustados** y el segundo **Índices Ajustados al Azar**.

5.3.1. Índices no ajustados

Este grupo está compuesto por cuatro índices, de los cuales **percent agreement** (a_o , pre 1901), **Holsti's CR** [183] y **Osgood's coefficient** [184] son matemáticamente equivalentes entre sí. En el desarrollo de estos índices, su diseño asume que todo el comportamiento de codificación es honesto y los acuerdos observados no contienen codificación aleatoria, por lo tanto, no hay necesidad de realizar ajustes por casualidad [152].

5.3.1.1. Percent agreement

El estadístico **percent agreement**, se utiliza con frecuencia para representar la confiabilidad porcentual en la medición entre examinadores, pero esta medida puede ser engañosa ya que no toma en cuenta **únicamente el acuerdo al azar entre examinadores**. La fórmula para su cálculo está dada por el número de casos en que los codificadores están de acuerdo (A) dividido por el número total de objetivos analizados (N). Krippendorff [185] and Neuendorf [181] denotaron al **percent agreement** como a_o [186], tal como se muestra en la Ecuación 2.

$$a_o = \frac{A}{N}$$

Ecuación 2. Percent agreement

En el desarrollo de su investigación Scott y su grupo de trabajo [187], observaron que el estadístico a_o era de uso común, quizás por ser tan común e intuitivo, sus primeros usuarios o críticos como Benini [188] no mencionaron quién lo inventó.

5.3.1.2. Holsti's CR

Holsti's y su grupo de trabajo, defendieron y utilizaron tanto el índice **percent agreement**, que muchos investigadores denominaron a este estadístico como el índice Holsti's CR [183].

Osgood's coefficient

Al igual que Holsti's, Osgood's y su grupo de trabajo defendieron y utilizaron tanto el índice **percent agreement**, que muchos investigadores denominaron a este estadístico como el índice Osgood's coefficient [184].

5.3.1.4. Rogot and Goldberg's A1

Rogot y Goldberg's, trabajaron en estudios de fiabilidad de métodos de clasificación. Los métodos de dichos estudios toman variedad de forma tales como:

Dos evaluadores, evaluadores o examinadores usando el mismo instrumento o procedimiento de clasificación.

Dos aplicaciones independientes de un determinado instrumento o procedimiento por un probador.

Dos instrumentos o procedimientos diseñados para producir el mismo resultado cuando se aplica al mismo individuo.

Aunque los métodos de fiabilidad que estudiaron difirieran entre ellos, partían del mismo diseño básico [189].

5.3.2. Índices ajustados al Azar

Los dieciocho índices de este grupo asumen que los codificadores maximizan deliberadamente la codificación aleatoria y limitan la codificación honesta a ocasiones dictadas por la casualidad, por lo que el acuerdo de probabilidad resultante debe estimarse y eliminarse, este grupo incluye tres subgrupos [152], los cuales se detallan a continuación:

5.3.2.1. Primer subgrupo de índices ajustados al Azar

Este primer subgrupo de nueve índices, estima el acuerdo de azar en función de la **categoría** en una escala de medición [152], los autores de estos índices son:

Gwet y su grupo de trabajo, presentaron en el año 2008 su trabajo por medio del artículo ***Computing inter-rater reliability and its variance in the presence of high agreement*** [190].

Grove y su grupo de trabajo, presentaron en el año 1981 su trabajo por medio del artículo ***Reliability Studies of Psychiatric Diagnosis: Theory and Practice*** [191].

Jason & Vegelius y su grupo de trabajo, presentaron en el año 1979 su trabajo por medio del artículo ***On generalizations of the G index and the phi coefficient to nominal scales*** [192].

Guilford y su grupo de trabajo, presentaron en el año 1961 su trabajo por medio del artículo ***Preparation of item scores for correlation between individuals in a Q factor análisis*** [193].

Holley & Guilford y su grupo de trabajo, presentaron en el año 1964 su trabajo por medio del artículo ***A note on the G index of agreement*** [194].

Perreault & Leigh y su grupo de trabajo, presentaron en el año 1989 su trabajo por medio del artículo ***Reliability of Nominal Data Based on Qualitative Judgments*** [195].

Brennan & Prediger y su grupo de trabajo, presentaron en el año 1981 su trabajo por medio del artículo ***Coefficient kappa: Some uses, misuses, and alternatives*** [196].

Byrt y su grupo de trabajo, presentaron en el año 1993 su trabajo por medio del artículo ***Bias, prevalence and kappa*** [197].

Potter & Levine-Donnerstein y su grupo de trabajo, presentaron en el año 1999 su trabajo por medio del artículo ***Rethinking validity and reliability in content análisis*** [198].

Guttman y su grupo de trabajo, presentaron en el año 1946 su trabajo por medio del artículo ***The test-retest reliability of qualitative data*** [199].

Maxwell y su grupo de trabajo, presentaron en el año 1977 su trabajo por medio del artículo ***Coefficients of agreement between observers and their interpretation*** [200].

Bennett y su grupo de trabajo, presentaron en el año 1954 su trabajo por medio del artículo ***Communications Through Limited Response Questioning*** [201].

5.3.2.2. Segundo subgrupo de índices ajustados al Azar

Este segundo subgrupo de ocho índices estima el acuerdo de azar en función de la ***distribución observada***, siendo la distribución el patrón por el cual los casos caen en categorías, que en ciertos casos puede ser extremadamente uniformes o extremadamente desiguales [152]. Los autores de estos índices son:

Krippendorff, y su grupo de trabajo, presentaron en los años 1970 y 1980 su trabajo por medio de los artículos ***Bivariate Agreement Coefficients for Reliability of Data*** [202], y ***Content Analysis: An Introduction to Its Methodology*** [185].

Rogot & Goldberg y su grupo de trabajo, presentaron en el año 1966 su trabajo por medio del artículo ***A proposed index for measuring agreement in test-retest studies*** [189].

Benini y su grupo de trabajo, presentaron en el año 1901 su trabajo por medio del artículo ***Principii di Demongraphia: Manuali Barbera Di Scienze Giuridiche Sociali e Politiche*** [188].

Byrt y su grupo de trabajo, presentaron en el año 1993 su trabajo por medio del artículo ***Bias, prevalence and kappa*** [197].

Cohen y su grupo de trabajo, presentaron en el año 1960 su trabajo por medio del artículo ***A Coefficient of Agreement for Nominal Scale*** [121].

Goodman & Kruskal y su grupo de trabajo, presentaron en el año 1954 su trabajo por medio del artículo ***Measures of Association for Cross Classifications*** [203].

Scott y su grupo de trabajo, presentaron en el año 1955 su trabajo por medio del artículo ***Reliability of Content Analysis: The Case of Nominal Scale Coding on JSTOR*** [187].

Siegel & Castellan y su grupo de trabajo, presentaron en el año 1988 su trabajo por medio del artículo ***Nonparametric Statistics for the Behavioral Sciences*** [204].

5.3.2.3. Tercer sub grupo de índices ajustados al Azar

Este tercer subgrupo tiene un solo índice que utiliza tanto la ***categoría como la distribución*** como factores principales [152]. Gwet y su grupo de trabajo, presentaron en los años 2008 y 2014 sus trabajos por medio de las publicaciones ***Computing interrater reliability and its variance in the presence of high agreement*** [190] y ***Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*** [205].

5.3.3. Selección del Índice estadístico del Corpus Life

Para el proceso de desarrollo del Corpus Life, el grupo **ASAPS** selecciono el estadístico Cohen's Kappa [121] para el proceso de medir la calidad de sus anotaciones, el cual utiliza una metodología de acuerdos que ajusta el efecto al azar entre evaluadores en la proporción de la concordancia observada para elementos cualitativos obtenidos de redes sociales y foros [10].

Los resultados de diferentes procesos de investigación que utilizaron el estadístico Cohen's Kappa, muestran que su uso conduce a clasificaciones con mayor grado de precisión respecto a otras medidas estadísticas descritas en la sección **índices estadísticos**, y, por lo tanto, conduce a soluciones de subconjuntos de características con características más relevantes [9]. Entre los proyectos de investigación que han utilizado el estadístico Cohen's Kappa, constan los siguientes:

O'Connor y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Promoting Reproducible Research for Characterizing Nonmedical Use of Medications Through Data Annotation: Description of a Twitter Corpus and Guidelines*** [102].

Sarker y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Portable automatic text classification for adverse drug reaction detection via multi-Corpus training*** [178].

O'Dea y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Detecting suicidality on Twitter*** [25].

Ben-David y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Comparison of classification accuracy using Cohen's Weighted Kappa*** [154].

Canales y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Exploiting a bootstrapping approach for automatic annotation of emotions in texts*** [76].

Fu y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Responses to a Self-Presented Suicide Attempt in Social Media*** [77].

Hallgren y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo ***Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial*** [78].

Tabachnick y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo *Experimental Designs Using ANOVA Epidemiology and Neurobehaviour of FASD in South Africa View project CyGaMEs View Project* [151].

Vieira y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo *Experimental Designs Using ANOVA Epidemiology and Neurobehaviour of FASD in South Africa View project CyGaMEs View Project* [80].

Vioules y su grupo de trabajo, quienes realizaron la publicación de su investigación por medio del artículo *Detection of suicide-related posts in Twitter data streams* [81].

Dada la importancia del estadístico Cohen's Kappa, la plataforma de software para el aprendizaje automático y la minería de datos WEKA¹³ lo utiliza como una de sus medidas de precisión [206].

En la construcción del Corpus del proyecto Life, se obtuvo un acuerdo moderado entre anotadores de 0,52 según la metodología del estadístico Cohen's Kappa, en el proceso de comparar 4 clases "No riesgo", "Posible", "Urgente" e "Inmediato". Sin embargo, se supone que, si el Corpus hubiera sido medido únicamente con dos clases Riesgo y Sin Riesgo, hubiera tenido una mayor concordancia. Este proceso de anotación es la base para entrenar nuevos datos con el fin de incrementar el tamaño del Corpus Life.

5.3.4. Software de procesamiento de texto

Por ser el proceso de anotación exhaustivo y consumir ingente cantidad de tiempo de anotadores e intérpretes, lo cual se traduce en un alto costo computacional, el poder seleccionar una herramienta de soporte al proceso de anotación adecuado, es crucial en orden de incrementar la calidad de las anotaciones y reducir costos [154]. La herramienta seleccionada para el proceso inicial de anotación del Corpus Life por parte del grupo de investigación ASAPS fue **GATE**¹⁴. El cual es un conjunto de herramientas de software de código abierto de soporte para el procesamiento de texto.

Con la herramienta **GATE**, se diseñó un método para obtener un Corpus objetivo, que respondiera al siguiente esquema:

1. Buscar semillas relacionadas con el suicidio en las redes sociales más comunes.

¹³ <https://www.weka.de/>

¹⁴ <https://gate.ac.uk/>

2. Obtener nuevas palabras después de haber localizado las semillas en Snippets, mensajes en el navegador de Twitter, navegador de Google, blogs y perfiles públicos etc.
3. Clasificar y almacenar la información de cada anotación de forma estructurada.

5.3.5. Esquema de anotación del Corpus Life

En el proceso de desarrollo de la plataforma Life, se desarrolló de forma supervisada el Corpus Life, el cual es parte medular de este proyecto, que inicialmente constaba de 102 textos en dos idiomas: inglés (71 textos) y español (31 textos), los cuales fueron recopilados de redes sociales como Facebook, Twitter, Instagram, blogs y foros. El Corpus tiene 2 clasificaciones diferentes e independientes: Nivel de alerta y Tipo de mensaje.

5.3.5.1. Alert Level

La clasificación Alert Level, consistió en anotar en el Corpus original los resultados de los análisis realizados por parte de los anotadores en base a 4 niveles de alerta Immediate, Urgent, Possible and No risk.

1. Immediate.

Immediate	
Descripción	Este es el nivel de alerta más alto, las personas incluidas en este grupo necesitan una derivación inmediata a los servicios de salud, ya que expresan claramente la idea de ser un suicida.
Detección de pistas	Descripción del comportamiento autolesivo con expresiones determinadas como ahora o un proceso complejo .
Ejemplos	<p>“In two hours, my life doesn’t make sense for anyone. I’ll be with God”</p> <p>“Mom, Dad, I love You and I’m sorry for everything’s, I’ll be better there rather than here. If you are reading this. It means that I won’t be with you”</p>

Fuente: Guía de anotación Life Corpus.

2. Urgent.

Urgent	
Descripción	El pensamiento suicida es notable, pero no de una manera tan inmediata. Los pensamientos depresivos repetitivos están relacionados con esta categoría.
Detección de pistas	Descripción del comportamiento autolesivo con expresiones como tenemos formas de suicidarnos ,

	pero además en los textos se observan textos de apoyo como parientes amados, siento que no soy nadie .
Ejemplos	"I don't find anything interesting so, I felt tired and exhausted. I don't want to continue anymore but I've got one child and he doesn't deserve that I'll be passed away".

Fuente: Guía de anotación Life Corpus.

3. Possible.

Possible	
Descripción	No todos los mensajes tienen el mismo riesgo. Los pensamientos indican episodios temporales de depresión, tristeza, insatisfacción con la vida, etc.
Detección de pistas	Posible, está relacionado con casos en los que tenemos dudas sobre si se tiene riesgo o no.
Ejemplos	"Life is a shit" "I have been looking for the truly meaning of file"

Fuente: Guía de anotación Life Corpus.

4. No risk.

No risk	
Descripción	Aquí se incluyen textos en los que el suicidio no se comenta en ningún parte, además hay pensamientos, mensajes y estímulos positivos.
Detección de pistas	Descripción de eventos positivos.
Ejemplos	"It's a beautiful day I can't stop myself from smiling" "The social networks are a good way to connect with the people".

Fuente: Guía de anotación Life Corpus.

5.3.5.2. Message Type

Para categorizar el tipo de textos del Corpus Life, se utilizaron dos propiedades "Primary Message Type" y "Secondary Message Type", ambos comparten la misma categoría, pero solo la primera es obligatoria. Específicamente, hay 12 categorías:

Auto-Pro-Suicide

Auto-Pro-Suicide	
Descripción	El comportamiento humano es complejo. Esa es la razón por la que una persona con pensamientos suicidas o con depresión se encasilla en esta categoría, también se incluyen sentimientos similares como el deseo de morir, pensamientos negativos (en general) y, por supuesto, toda la gente que ha hablado de intentos de suicidio.
Detección de pistas	Hechos sobre la depresión que coinciden con los comportamientos " Auto pro-suicidas ", están incluidos

	aquí, por lo que no se debe marcar en la propiedad TypeMessage de la categoría depresión, debe marcarse como Urgente o Inmediato. En caso de acuerdo con el tema de la muerte, pero no se menciona como posibilidad, se puede marcar en el TypeMessage Possible . Los comportamientos autolesivos con el fin de revelar una expresión concreta de enojado, pero no relacionado con el suicidio no se incluyen en esta categoría.
Ejemplos	I want to kill myself

Fuente: Guía de anotación Life Corpus.

No-Auto-ProSuicide

No-Auto-ProSuicide	
Descripción	No todo el mundo está de acuerdo con la idea del suicidio, por lo que en esta categoría se incluyen estas personas. El movimiento de suicidología, pensamientos positivos, textos de apoyo y personas que habían superado los pensamientos suicidas y el trastorno depresivo.
Detección de pistas	
Ejemplos	“I understand you, but life is beautiful so, if you need talk, I’m here for you”. “Few months ago, I was in your situation and thought about suicide”.

Fuente: Guía de anotación Life Corpus.

Irony.

Irony	
Descripción	La ironía consiste en un uso figurado del lenguaje. Esa es la razón por la que utilizaron una operación cognitiva en un contexto adecuado. Por ese objetivo, el receptor tiene que hacer una comparación y contraste de los textos.
Detección de pistas	
Ejemplos	“It’s like choosing soup over the salad”

Fuente: Guía de anotación Life Corpus.

Mental Illness or Disorder.

Mental Illness or Disorder	
Descripción	Aquí se incluyen las manifestaciones de enfermedades o trastornos mentales.

Detección de pistas	Solo se incluyen enfermedades mentales diagnosticadas previamente.
Ejemplos	"I've been taken pills for major depression diagnosed few years ago".

Fuente: Guía de anotación Life Corpus.

Suicide Curiosity.

Suicide Curiosity	
Descripción	Las personas incluidas en esta categoría expresan una especie de interés por el tema del suicidio.
Detección de pistas	
Ejemplos	"I don't know why but I feel a bit curiosity about the idea of suicide, If I had all the tools needed, I would try"

Fuente: Guía de anotación Life Corpus.

Euthanasia.

Euthanasia	
Descripción	Situaciones en las que el suicidio asistido es una opción, según el estado de un problema de salud, que muestra enfermedades terminales o extrañas. Esta categoría se creó para incluir ese tipo de casos.
Detección de pistas	
Ejemplos	"I agree with Euthanasia".

Fuente: Guía de anotación Life Corpus.

Mysticity.

Mysticity	
Descripción	Escribirlo puede ser un riesgo o un factor de protección, pero para nuestro objetivo (tener un Corpus representativo del suicidio) este tipo de historias puede ser una forma de expresar cosas que de otras formas no podrían ser posibles. Además, en esta categoría se incluyen factores religiosos.
Detección de pistas	El tipo de textos incluidos en esta categoría expresan, de manera fantástica, la idea de morir por no mencionarlo específicamente. En esta categoría, no tenemos que marcar la segunda propiedad Secondary Message Type.
Ejemplos	"Only I'll be happy When I lived with God" "She was like an angel and our soul were in touch in that night"

Fuente: Guía de anotación Life Corpus.

Undefined.

Undefined	
Descripción	En este apartado pondremos los casos que no se pueden definir en las otras categorías.
Detección de pistas	Si el mensaje no es ni suicida ni depresivo, el nivel de alerta no es Riesgo.
Ejemplos	"The only solution is to close the eyes I have a problem"

Fuente: Guía de anotación Life Corpus.

Instigator.

Instigator	
Descripción	Esta categoría incluye a las personas que animan a otras personas a suicidarse.
Detección de pistas	Las formas imperativas son muy comunes en esta categoría.
Ejemplos	"Try to disappear of this ruin world". "The life is so bored that It doesn't make sense being alive".

Fuente: Guía de anotación Life Corpus.

Sadness/Melancholy.

Sadness/Melancholy	
Descripción	Tristeza que dura (depende de la persona) un período indeterminado de tiempo, a menudo sin ninguna razón obvia.
Detección de pistas	Los textos poéticos y sin contexto son adecuados aquí. Además, podemos incluir textos en los que las palabras claves no pertenecen a otra categoría. Entendemos que los mensajes incluidos aquí solo los tienen en la propiedad AlertLevel, la categoría Sin riesgo.
Ejemplos	"All finished without a beginning". "I'm broken inside and empty. I need a coffee in a rainy day, hearing the sound of my soul".

Fuente: Guía de anotación Life Corpus.

Citation.

Citation and Depression	
Descripción	Cuando expresas algo en tus propios términos, le haces saber a la gente con el tipo de mensaje que puede transmitir, pero si no utiliza sus palabras, desea expresar la misma idea, pero el lenguaje utilizado no es lo mismo, entonces, no es su forma natural de expresar ideas.

Detección de pistas	Uso de comillas simples / dobles. En esta categoría, no se tiene que marcar la segunda propiedad SecondaryMessageType.
Ejemplos	... “ That which does not kill us makes us stronger” (Friedrich Nietzsche).

Fuente: Guía de anotación Life Corpus.

Depression.

Citation and Depression	
Descripción	Depresión Puede ser un estado mental o un trastorno particular que combina comportamientos como fatiga, falta de sueño, trastornos alimentarios, falta de concentración, etc.
Detección de pistas	Las palabras como todo o nada tienen sentido aquí si el contexto combina con el tema. Vacío indica algo que dura mucho tiempo. Aquí no incluimos la depresión que había sido previamente diagnosticado. Siempre, se seleccionará la propiedad AlertLevel, como mínimo, con la Categoría posible si marcamos esta categoría de depresión.
Ejemplos	“I don't want to eat, I can't sleep, I don't want to do anything” “no good for nothing”. “I don't find rest during the night, if only I could sleep well”.

Fuente: Guía de anotación Life Corpus.

Una vez concluido el proceso de anotación se obtuvo un Corpus de suicidio con 102 muestras en sus diferentes niveles de alerta. Aunque la principal ventaja del Corpus es que su calidad ha sido evaluada, existe el inconveniente de que sus categorías cuentan con pocas o ninguna muestra y su desequilibrio es notorio. Por ello luego de realizar experimentos iniciales no tan satisfactorios, se decidió utilizar únicamente la clasificación de nivel de alerta, ya que esto reduce el número de categorías y proporciona una mayor cantidad de muestras por tipo para lograr una mayor significancia estadística. Tal como se muestra en la Tabla 30.

Nivel de alerta	Cantidad	EN	ES
No riesgo	70 (68,6%)	45	25
Urgente	19 (18,6%)	15	4
Posible	8 (7,8%)	6	2
Inmediato	5 (4,9%)	5	0

Tabla 30. Numero de ejemplos por cada tipo de alerta.

A pesar de haber decidido utilizar únicamente el nivel de alerta del Corpus Life, el desbalance de este en las clases que indican algún tipo de riesgo (Tabla 30), puede generar resultados no confiables al realizar la evaluación con los algoritmos de clasificación. Por lo tanto, para optimizar el resultado de los experimentos, se decidió

generar un nuevo Corpus con solo dos clases: Riesgo y No Riesgo. Así la clase Riesgo contendrá todos los mensajes de alerta que indicaran algún tipo riesgo (Urgente, Posible e Inmediato) con 32 muestras y dejar la clase No Riesgo con 70 muestras, como se muestra en la Tabla 31.

Nivel de alerta	Cantidad	EN	ES
No riesgo	70 (68.63%)	45	25
Riesgo	32 (31.37%)	26	6

Tabla 31. Numero de ejemplos por tipo de alerta riesgo y No riesgo.

Con el Corpus Life con dos clases, se han realizado una serie de experimentos, varios de los cuales se han sido publicados en forma de artículos, entre los cuales se tiene **“Assessment of supervised classifiers for the task of detecting messages with suicidal ideation”** [112], y **“An Unsupervised Learning Approach to Automatically Categorize Potential Suicide Messages in Social Media”** [173].

En el desarrollo del primer grupo de experimentos [112], se llevó a cabo por medio un análisis sistemático de todas las posibles características textuales del corpus Life, con 28 algoritmos de clasificación contenidos en la plataforma para el aprendizaje automático y la minería de datos Weka [207]. El cual concluyo que era necesario incrementar el número de registros del corpus para alcanzar un mejor rendimiento, a más de probar nuevos algoritmos de clasificación. El desarrollo de estos experimentos consta en el subcapítulo 7.1 **Experimentos 1**, de este documento.

Para incrementar el número de registros del Corpus Life, se desarrolló un método automático semi-supervisado, por medio del cual detectar, clasificar y anotar textos extraídos de entornos digitales relacionados con el suicidio. Siendo en este proceso necesario desarrollar un método que basado en las muestras anotadas de forma supervisada en el Corpus Life, **Realizar la validación de la calidad de los textos relacionados con el suicidio obtenidos desde entornos digitales**. El desarrollo de estos experimentos consta en el subcapítulo 7.2 **Experimentos 2**, de este documento y han sido publicados por medio del artículo “Bootstrapping semi-supervised annotation method for potential suicidal messages” [208].

Capítulo VI

6. Método de traducción paralela

En este capítulo se desarrollará el **tercer objetivo específico** “**Generar Corpus paralelos a diferentes idiomas a partir del Corpus de suicidio escrito en idioma inglés**”. Para esto analizaremos los métodos actuales de traducción automática de idiomas.

6.1. Introducción

En un mundo globalizado, en el cual los eventos que acontecen en cualquier parte del mundo son transmitidos por medio de la televisión o redes sociales, en donde plataformas en línea nos ofrecen acceso a catálogos de series, películas, cursos de capacitación [209][210][211], compras etc. La traducción automática de contenido es una necesidad para satisfacer a sus usuarios y ampliar sus campos de negocios [212].

Los 44 zettabytes de información disponible actualmente en el mundo, han sido generados entre los aproximadamente 7.139 idiomas en el mundo. Entre los idiomas con mayor número de hablantes y mayor incidencia en la generación de información en el mundo, se tiene al inglés con 1.348 millones de hablantes, Chino Mandarín con 1120 millones, hindú con 600 millones, español con 543 millones, árabe con 274 millones, bengalí con 268 millones, francés con 267 millones, ruso con 258 millones, portugués con 258 millones y el urdu con 230 millones [213].

En el proceso de generar información, las plataformas de redes sociales juegan un papel clave ya que posibilitan que las personas a través de ellas puedan expresar sus sentimientos [63][64], [65][65][66]. Estos sentimientos expresados en las redes sociales, representan una oportunidad para especialistas en las diferentes áreas del conocimiento [9], a través del *aprendizaje automático o Machine Learning*, para descubrir patrones, tendencias y relaciones entre datos, que les permita alcanzar un conocimiento para predecir resultados [61] y alcanzar una ventaja competitiva sobre los demás [9].

Dado el gran número de idiomas en que se puede generar información de un tema específico de forma escrita u hablada, diversos grupos de investigadores se han visto en la necesidad de realizar experimentos no solo en el idioma original de la investigación, sino en múltiples idiomas, en muchos casos por medio de la traducción paralela de Corpus lingüísticos [125][214].

Los grupos de investigación que utilizan Corpus Lingüísticos para sus procesos de investigación, en más de un idioma, utilizan por lo general para su traducción Corpus paralelos, a partir de las anotaciones semánticas del Corpus base [215][216].

6.2. Proyectos realizados con corpus paralelos

Dado que la mayor parte de los Corpus desarrollados hasta la actualidad han sido anotados con datos en idioma inglés, contar con un método de traducción en forma paralela a diferentes idiomas es necesario [217], siempre, que se quiera realizar

experimentos de un mismo tema de investigación en diferentes idiomas. Entre los diferentes grupos de investigación que trabajan con Corpus paralelos en diferentes áreas del conocimiento tenemos:

Zhu y su equipo, centraron su investigación en el desarrollo de una metodología que genere una red neuronal de atención jerárquica paralela de nivel, que les permitió codificar oraciones monolingües, con el objeto de construir un clasificador para extraer oraciones paralelas.

Rapp y su equipo, propuso utilizar un enfoque híbrido que empareja la inserción de oraciones multilingües y un clasificador supervisado para identificar pares de oraciones paralelas, para ello utilizaron características de signos de puntuación de origen-destino y características morfosintácticas para construir un clasificador SVM binario [218].

Jolly y su equipo, trabajaron en la creación y evaluación de un Corpus paralelo para traducción automática del idioma inglés al idioma Penjabi el cual es un idioma de bajos recursos, utilizando para ello métricas de permutación de alineación de oraciones [219].

Resnik y su equipo, trabajaron en el sistema STRAND, con el objeto de extraer Corpus paralelos en la Word Wide Web a gran escala, para lo cual utilizaron aprendizaje supervisado basado en características estructurales de los documentos para mejorar el desempeño de la clasificación [147].

Koehn y su equipo, trabajaron a partir del Corpus de textos paralelo del parlamento europeo, el cual contiene los textos de las actas generadas y traducidas paralelamente a 11 idiomas. Para la construcción de un sistema estadístico de traducción automática SMT para 110 pares de idiomas [126].

Como se puede denotar en las conclusiones de los artículos científicos generados por los proyectos de investigación expuestos, la tecnología de traducción paralela no ha alcanzado una madurez total. Pero sus avances son significativos, y mejoraran muchos más, gracias a que existen grupos de investigación interesados en mejorar los métodos actuales de traducción. Dada la necesidad de analizar la información generada por las personas en los más de 7000 idiomas existentes en nuestro mundo y poder satisfacer las necesidades de investigación, en todo tipo de campos del conocimiento.

6.3. Traducción paralela del Corpus de suicidio

La plataforma Life, tiene como objeto detectar de forma automática textos de usuarios que posten en redes sociales mensajes con sentimientos suicidas, incitación al suicidio o síntomas depresivos, en cualquier lugar u idioma. Para con la ayuda de instituciones especializadas, intentar contactar a estos usuarios, persuadirlos que desista de su intención suicida, brindándoles ayuda de forma urgente, gratuita, anónima y especializada en la crisis emocional que estén viviendo [139].

Para cumplir con los objetivos de esta plataforma, el Corpus Life descrito en la subsección 4.3.5. *Esquema de anotación del corpus Life*, se debe escoger un método de traducción paralela que permita traducir sus textos originales a otros idiomas. Para que estos textos no solo estén escritos en un idioma determinado, sino que estén escritas en la mayor cantidad de idiomas posibles. Lo que conlleva, que, a partir del Corpus original, se creen Corpus paralelos en diferentes idiomas, con el fin de que se

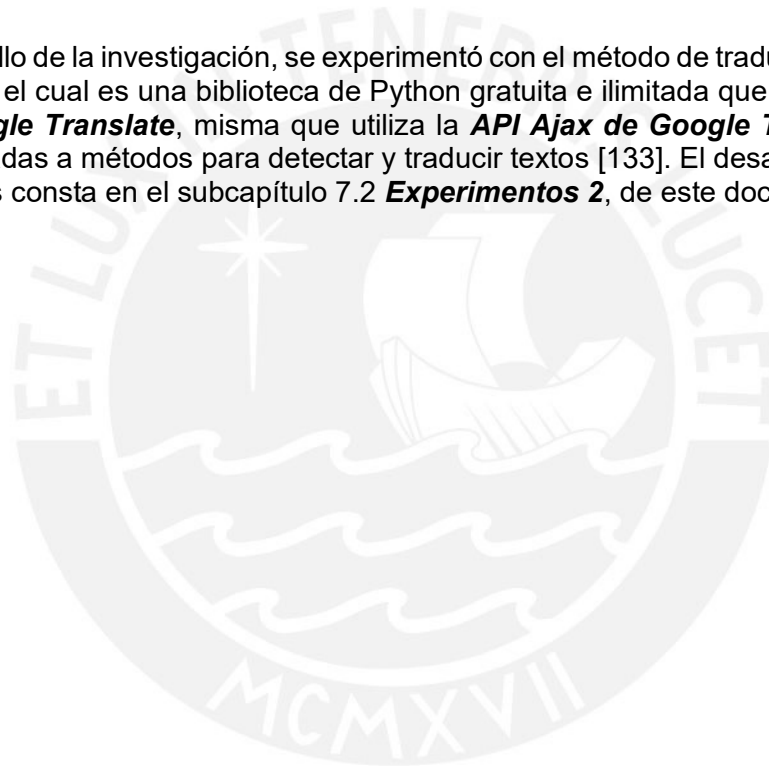
pueda intentar prevenir suicidios de usuarios que postean sus intenciones suicidas a través de las redes sociales en cualquier lugar o idioma.

Este proceso de traducción automática se inicia con la traducción al idioma inglés de los 31 textos del Corpus Life en idioma español al idioma inglés, con el propósito de que, con un mayor número de muestras en un mismo idioma, tener mejores resultados al momento de realizar experimentos para intentar detectar usuarios suicidas en redes sociales (Tabla 32).

NIVEL DE ALERTA	CANTIDAD	%
No Riesgo	70	63,4%
Riesgo	32	21,1%

Tabla 32 Textos en idioma inglés, por nivel de alerta

En el desarrollo de la investigación, se experimentó con el método de traducción paralela Googletrans, el cual es una biblioteca de Python gratuita e ilimitada que implementa la **API de Google Translate**, misma que utiliza la **API Ajax de Google Translate** para realizar llamadas a métodos para detectar y traducir textos [133]. El desarrollo de estos experimentos consta en el subcapítulo 7.2 **Experimentos 2**, de este documento.



Capítulo VII

7. Análisis de resultados y validación de la propuesta

Es este capítulo, se describen dos grupos de experimentos, efectuados en el marco de la actual investigación doctoral.

7.1. Experimentación 1

Para este primer grupo de experimentos, utilizamos técnicas de aprendizaje automático para llevar a cabo un análisis sistemático de todas las combinaciones posibles de características textuales del Corpus Life, con 28 algoritmos de clasificación nativos de la plataforma de software Weka con sus parámetros por defectos [207]. Con la finalidad de determinar cual algoritmo de clasificación y que características del Corpus, son las más adecuadas para el proceso de clasificación automática de mensajes que contengan ideación suicida, una vez que estos sean ubicados y extraídos desde las redes sociales [112].

7.1.1. Evaluación de clasificadores supervisados para la tarea de detección de mensajes con ideación suicida

Previo a iniciar los experimentos, el Corpus Life fue preprocesado para extraer diferentes tipos de características utilizando técnicas de PLN y probándolas en diferentes combinaciones para descubrir qué características eran más adecuadas para la clasificación automática de mensajes con ideación suicida. Las características utilizadas fueron:

Bag of Words (WORD). Bolsa de palabras del documento sin ninguna modificación.

Bag of Stems (STEM). Las palabras del documento mediante la realización de un proceso de derivación para reducir la variedad de derivaciones de una sola palabra.

Bag of Lemmas (LEMMA). En lugar de usar cada palabra, su lema se utilizó con el mismo objetivo que el uso de la raíz, así como para averiguar qué técnica (derivación o lematización) era mejor.

Bag of SYNSETS (SYNSET). El SYNSET de WordNet de cada término utilizado en el Corpus se obtuvo para ampliar la cobertura del Corpus eligiendo palabras con el mismo significado en lugar de la palabra literal. Estas características se extrajeron utilizando *Freeling 4.0* [220].

Bag of POS (POS). También se obtuvo la parte del discurso de cada término para utilizarlo como características. En experimentos previos se ha podido

comprobar que el uso de esta información léxica puede ayudar a discriminar las notas de suicidio auténticas de las falsas debido a que las primeras suelen hacer más uso de, por ejemplo, adjetivos entre otras características.

Característica	Con números		Sin números	
	Con stopwords	Sin stopwords	Con stopwords	Sin stopwords
WORD	1713	1474	1705	1466
STEM	1434	1261	1433	1254
LEMMA	1430	1260	1422	1252
SYNSET	4885	4507	4885	4507
POS	173	148	173	148

Tabla 33 Numero de características por cada tipo de característica, con y sin stopwords.

Con las características expuestas en la tabla 33, se lanzó un experimento por cada combinación, por ejemplo, bag of words, bag of words y POS, bag of words, POS y SYNSETS, etc. Independientemente de la combinación utilizada, se repitió el experimento con y sin stopword y, además, manteniendo los valores numéricos o reemplazándolos por una etiqueta única `_NUM_`. Estas dos técnicas se utilizaron para ver si reduciendo el número de características del Corpus se podía obtener los mismos o mejores resultados.

Por cada una de estas combinaciones, se realizaron experimentos con palabras vacías y sin ellas o con la etiqueta del número sustituto. Un total de 124 experimentos con distintas combinaciones de características se lanzó por cada uno de los 28 algoritmos de aprendizaje automático utilizados de los 45 que se encuentran de forma nativa en la plataforma de software para el aprendizaje automático Weka [207]. Cabe mencionar que los 17 restantes algoritmos no se los utilizo por cuanto no contaban con parámetros suficientes o por su elevado tiempo de procesamiento. Estos primeros experimentos fueron desarrollados en el intérprete de comandos de Linux.

7.1.2. Enfoques de aprendizaje automático

En este apartado se da una descripción general de la metodología empleada para determinar el clasificador supervisado y las características más adecuadas para la clasificación automática de mensajes con ideación suicida mediante el Corpus Life. Los experimentos se llevaron a cabo extrayendo ciertas características del Corpus como se explica en la sección 7.1.1, utilizando 124 combinaciones diferentes. Es decir, con las características extraídas del Corpus: WORD, STEM, LEMMA, SYNSET y POS, se realizaron experimentos para cada una de estas características de forma individual.

Para realizar la experimentación se utilizó el software Weka [206][207]. Para cada una de las combinaciones descritas se entrenó un modelo con cada uno de los 28 algoritmos clasificadores que se muestran en la Tabla 34, junto con los parámetros utilizados con el fin de facilitar la reproducción de estos. Al final, se realizaron 3.472 experimentos, combinando las características con cada uno de los algoritmos clasificadores (124 combinaciones de características y 28 clasificadores).

Para evaluar la significación estadística de los resultados, repetimos cada experimento con 30 divisiones aleatorias diferentes del Corpus para realizar una prueba estadística ANOVA. En estos experimentos, usamos los valores dados por el software Weka definidos por las ecuaciones: F-measure (ecuación 3), ROC Área (ecuación 4), Precision (Ecuación 5) y Recall (Ecuación 6).

$$F_1 = 2 \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

Ecuación 3. F-measure

$$\text{ROC Area: } TPRate = \frac{TP}{TP + FN}, FPRATE = \frac{FP}{TN + FP}$$

Ecuación 4. ROC Area

$$\text{precision} = \frac{TP}{TP + FP}$$

Ecuación 5. Precision

$$\text{recall} = \frac{TP}{TP + FN}$$

Ecuación 6. Recall

Si bien estas medidas son las más utilizadas en la evaluación de clasificadores, tenemos algunos problemas para reconocer el modelo más adecuado para las tareas de detección de suicidios. Seguidamente en la tabla 34 se muestran los algoritmos de clasificación con los parámetros por defectos utilizados en este experimento.

Algoritmos de clasificación	Parámetros por defecto
BayesNet	-D -Q weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- - A 0.5 \ SimpleLogistic & -I 0 -M 500 -H 50 -W 0.0
SMO	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -C 250007 -E 1.0" IBK
IBK	-K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\" KSTAR
KStar	-B 20 -M a
AdaBostM1	-P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump

	ATTCLASS
Baggind	-P 100 -S 1 -num-slots 1 -I 10 -W weka.classifiers.trees.REPTree -- -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0
CVParamSelection	-X 10 -S 1 -W weka.classifiers.rules.ZeroR
MultiClassifier	-M 0 -R 2.0 -S 1 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M -1 -num-decimal-places 4
MultiClassUp	sin parámetros
MultiSchema	-X 0 -S 1 -B weka.classifiers.rules.ZeroR
RandomCommittee	-S 1 -num-slots 1 -I 10 -W weka.classifiers.trees.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1
RandomFiltClass	-S 1 -F weka.filters.unsupervised.attribute.RandomProjection -W weka.classifiers.lazy.IBk -- -K 1 -W 0 -A weka.core.neighboursearch.LinearNNSearch RANDOMSUB
RandomSubSpace	-P 0.5 -S 1 -num-slots 1 -I 10 -W weka.classifiers.trees.REPTree -- -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0 STACKING
Stacking	-X 10 -M weka.classifiers.rules.ZeroR -S 1 -num-slots 1 -B weka.classifiers.rules.ZeroR VOTE
Vote	-S 1 -B weka.classifiers.rules.ZeroR -R AVG
WeightedDistances	-S 1 -W weka.classifiers.rules.ZeroR
InputMappedClassifier	-I -trim -W weka.classifiers.rules.ZeroR
DecisionTable	-X 1 -S weka.attributeSelection.BestFirst
JRip	-F 3 -N 2.0 -O 2 -S 1
OneR	-B 6
PART	-M 2 -C 0.25 -Q 1
ZeroR	-output-debug-info
HoeffdingTree	-L 2 -S 1 -E 1.0E-7 -H 0.05 -M 0.01 -G 200.0 -N 0.0
J48	-R -N 3 -Q 1 -M 2
LMT	I -1 -M 15 -W 0.0
RandomForest	-K 0 -M 1.0 -V 0.001 -S 1
RandomTreede	M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0

Tabla 34 Algoritmos y parámetros por defectos utilizados in los experimentos.

El fin de estos experimentos, es detectar la mayor cantidad de casos de ideación suicida. Intentando minimizar los falsos positivos para no sobrecargar los servicios de urgencias con falsas alarmas. Pero maximizando los verdaderos positivos de la clase que supongan algún riesgo. Se puede determinar, que ninguna métrica refleja ese valor, ni se cree que exista un valor adecuado, más bien depende de las preferencias de los trabajadores del servicio de emergencia o de cómo se priorizan los mensajes. Por ejemplo, si en un servicio de prevención de suicidios generalmente hay pocas alertas, podría ser interesante ampliar la cobertura hasta que se generen un mayor número de alertas, o en cambio, puede interesar asegurar que los mensajes que llegan sean en su mayoría verdaderos positivos, reduciendo así las falsas alarmas [221].

En ambos extremos, los algoritmos clasificadores o las características utilizadas serían diferentes e inicialmente, no podríamos descartar ninguna. Por este motivo se decidió representar los clasificadores con métricas bidimensionales, donde el eje X que representa la proporción de aciertos en las clases que representan algún nivel de riesgo

dividido por las muestras de riesgo total, como aparece en la Ecuación 7, y en el eje Y, la proporción de aciertos en las clases que representan poco o ningún riesgo en comparación con el total como se muestra en la Ecuación 8.

$$precision_{risk} = \frac{TP_{risk}}{TP_{risk} + FP_{risk}}$$

Ecuación 7. Precision_risk

$$precision_{no_risk} = \frac{TP_{no_risk}}{TP_{no_risk} + FP_{no_risk}}$$

Ecuación 8. Precision_No_risk

De esta forma, se representó en un solo gráfico (Ilustración 10) los resultados de los clasificadores, en el cual los puntos que más interesen en primera instancia serán los que se representen más cerca del valor 1 de estas dos métricas. Sin embargo, y como se verá, cuanto más se intente aumentar el primer valor más tenderá a reducir el segundo y viceversa. Así, en función de las necesidades, se tendrá que elegir los algoritmos que más interesen.

Una vez obtenidos los resultados de los experimentos realizados con Weka, estos fueron procesados por medio del estadístico Chi² para verificar cuales eran estadísticamente normales y cuáles no. Ecuación 9.

$$x^2 = \sum_{i=1}^k \left(\frac{fo_i - fe_i}{fe_i} \right)^2 > X^2$$

Ecuación 9. Estadístico Chi2,

Donde fo_i es la frecuencia observada y fe_i es la frecuencia esperada.

A más de discriminar los experimentos con valores estadísticamente anormales, el estadístico Chi₂ permitió discriminar aquellos experimentos en los cuales la desviación típica es igual a cero (clasificadores que independientemente de la muestra de entradas, siempre clasificaban como "No Risk", es decir de la clase con más muestras), como resultado se obtuvo que 1.216 experimentos eran viables de acuerdo con la formula planteada y son los que tomamos en cuenta para continuar con el proceso.

Inmediatamente se ejecutó el estadístico ANOVA para comparar semejanzas y diferencias estadísticas entre los 1.216 experimentos resultantes del proceso estadístico Chi².

$$SS_{TOTAL} = \sum X_i^2 - \frac{T^2}{N} \quad T = \text{suma total de tiempos}$$

Ecuación 10. SS_total

$$SS_{TRAT} = \sum T_i^2 - \frac{T^2}{N} \quad T_i = \text{suma por tratamiento}$$

Ecuación 11. SS_trat

$$SS_{ERROR} = SS_{TOTAL} - SS_{TRAT}$$

Ecuación 12. SS_error

$$F = \frac{\frac{SS_{TRAT}}{a-1}}{\frac{SS_{ERROR}}{N-a}}$$

Ecuación 13. F

a: Número de clasificadores

N: Número de veces

Como resultado de realizar una comparación de todos contra todos, entre los experimentos estadísticamente significativos, utilizando el estadístico ANOVA, se obtuvo 739.935 comparaciones, de las cuales 71.983 fueron estadísticamente significativas.

7.1.3. Resultados

Para obtener los resultados, se realizaron 3.472 experimentos con diferentes combinaciones de características y algoritmos clasificadores, como se describe en la Sección 7.1.2. Se utilizó el Corpus Life, con las 4 clases originales ("No riesgo", "Posible", "Urgente" e "Inmediato"), luego de lo cual se agruparon las características en solo dos ("No riesgo " y "Riesgo "), donde las 2 clases que indican algún riesgo de ideación suicida o depresión se han consolidado en una. Seguidamente se realizó un análisis estadístico de los resultados mediante el estadístico ANOVA con un margen de confianza de $p < 0,05$. En la Tabla 35 se muestran los 20 mejores resultados de la medida F-measure y del Área ROC para la versión del Corpus Life con cuatro clases.

Características del Corpus	Clasificador	F-measure	ROC Área
(1) POS,NUM (2) POS,SYNSET,NUM	KStar	0,715	0,749
(3) POS,SYNSET (4) POS	KStar	0,715	0,749

(5) POS,SYNSET,NUM (6) POS,SYNSET	RandomCommittee	0,704	0,718
(7) POS,STEM,SYNSET,LEMMA,WORD,NUM	RandomCommittee	0,703	0,718
(8) STOPWORD,SYNSETS,LEMMA	RandomCommittee	0,702	0,718
(9) POS,STEM,WORD,NUM	RandomCommittee	0,701	0,718
(10) STEM,SYNSETS (11) STEM	KStar	0,700	0,724
(12) STOPWORD,STEM,SYNSETS,LEMMA,WORD	KStar	0,700	0,724
(13) WORD,NUM (14) STOPWORD,LEMMA,WORD	KStar	0,700	0,724
(15) STOPWORDS,STEM,LEMMA,WORD	KStar	0,700	0,724
(16) STOPWORDS,STEM,LEMMA,WORD,NUM	RandomCommittee	0,699	0,704
(17) STOPWORDS,STEM,SYNSETS,LEMMA	RandomCommittee	0,699	0,704
(18) WORD (19) POS,STEM,LEMMA,WORD	RandomCommittee	0,699	0,704
(20) LEMMA,WORD,NUM	SMO	0,699	*
POS: Part of Speech; SYNSET: Wordnet Synsets; WORD: Bag of Words; STEM: Stem of words; LEMMA: Lemma of words; NUM: without numbers; STOPWORD: without stopwords			
Nota: Cada combinación de funciones de entrenamiento está numerada por (#), (*) este valor no es estadísticamente significativo.			

Tabla 35 Resultados de los principales valores de la evaluación del Corpus con diferentes características y diferentes clasificadores para la versión del Corpus con 4 clases.

De acuerdo con los resultados presentados en la Tabla 35, se determinó que el clasificador KStar fue el de mejor desempeño en todos los experimentos, seguido por los clasificadores RandomCommittee y SMO. También se observó que el enfoque de aprendizaje automático con mejor desempeño fue el clasificador KStar, que logró una mayor reducción de funcionalidades: POS (Part of Speech), NUM (sustitución de los diferentes números por un solo teclado NUM y el SYNSET de WordNet que representa la categoría semántica del término).

Sin embargo, se percibió que los POS con diferentes términos en los textos lograban los mismos resultados y que con la combinación con otras características se volvían más complejos. Esto parece indicar que, para un Corpus sobre ideación suicida de este tamaño, la categoría léxica es más que suficiente para obtener los mejores resultados. Esto determina que el estado de ánimo puede manifestarse en las características lingüísticas que se utilizan para comunicarse [117].

Otro clasificador que dio resultados similares, pero significativamente más bajos fue el RandomCommittee. En éste se encontró el mismo patrón, donde POS fue la característica predominante en combinación con otras, no habiendo nada estadísticamente significativo para diferenciarlas. Se encontraron diferencias mínimas con las características de STEM y WORD, que representaban bolsas de palabras, la primera con lematizador y la segunda sin lematizador, ya sea que se eliminaran palabras vacías

o no, o cuando se le agregaban nuevas características como el SYNSET o el lemma de las palabras. Las otras combinaciones de características y clasificadores produjeron valores significativamente más bajos.

Debido al pequeño tamaño del Corpus, había clases en las que el número de muestras era muy bajo, y era difícil crear un modelo de aprendizaje que fuera capaz de clasificar correctamente las muestras de esa clase. Esto podría provocar resultados insignificantes para esas clases, no sean significativas y generaran ruido en la evaluación. Por esto se decidió fusionar las 3 clases de riesgo (Posible, Urgente e Inmediato) en una sola. Y llegar a un Corpus con dos clases: El primero de mensajes sin riesgo de suicidio con 70 muestras (Sin riesgo) y el segundo con mensajes con algún nivel de riesgo con 32 textos (Riesgo). Los resultados de esta evaluación se pueden ver en Tabla 35.

Características del Corpus	Clasificador	F1	ROC Área
(1) POS,SYNSETS,NUM	KStar	0,746	0,810
(2) POS,SYNSETS,lemma,Word	RandomCommitte	0,716	0,768
(3) stopwords,POS,stem,NUM	RandomizableFiltered	0,715	0,675
(4) stopwords,POS,stem,SYNSETS, lemma,word,NUM	RandomizableFiltered	0,713	0,671
(5) POS,NUM	RandomTree	0,715	0,668
(6) stopwords,POS,stem	RandomizableFiltered	0,712	0,663
(7) stopwords,POS,stem,SYNSETS, word,NUM	RandomTree	0,714	0,663
(8) stem,lemma,NUM	RandomizableFiltered	0,709	0,662
(9) stopwords,POS,stem,word	RandomizableFiltered	0,699	0,662
(10) stopwords,POS,SYNSETS, lemma,word,NUM	RandomizableFiltered	0,700	0,661
(11) POS,SYNSET,word,NUM	RandomizableFiltered	0,706	0,600
(12) POS,stem,NUM	RandomTree	0,704	0,655
(13) stopwords,stem	RandomizableFiltered	0,689	0,654
(14) POS,stem	RandomTree	0,700	0,650
(15) POS,stem,SYNSETS,NUM	RandomTree	0,685	0,643

Tabla 36 Resultados de los principales valores de la evaluación con diferente clasificador y diferentes categorías para la versión del Corpus con 2 clases ("Riesgo" y "No Riesgo").

Como se esperaba, los resultados con 2 clases superaron significativamente los resultados evaluados con 4 clases, logrando una mejora de aproximadamente 7 puntos sobre el Área ROC en el mejor clasificador. Además, se observa el mismo patrón KStar con las características más simples de la publicación junto con SYNSETS y NUM, dio los mejores resultados. Cabe señalar que los resultados de los algoritmos que utilizan Decision Tree (RandomCommittee, RandomizableFiltered y RandomTree) se colocaron en casi todas las primeras posiciones, excepto en un caso. En la mayoría de ellos, las características POS estaban presentes.

El Life Corpus utilizado en el presente estudio fue creado para ser utilizado como base en un sistema de prevención de suicidios basado en el aprendizaje automático. Este sistema (aún en desarrollo) parte de la detección de mensajes de posibles ideas suicidas para enviar avisos a las agencias de prevención del suicidio. Por lo tanto, requiere un

equilibrio entre verdaderos positivos y falsos positivos de la clase de riesgo con verdaderos positivos y falsos positivos de la clase que no está en riesgo. Esto se debe a que el sistema debe detectar los mensajes de ideación suicida como sea posible para llegar a la mayor cantidad de personas que necesitan ayuda, pero reducir los avisos de falsos positivos y evitar sobrecargar los servicios de prevención.

El equilibrio depende en gran medida de la capacidad de estas agencias de prevención [222][139] y puede variar según los recursos disponibles. Por tanto, en sistemas desequilibrados, medidas como F-measure o el Área ROC pueden no ser las más adecuadas para detectar qué clasificador es el mejor, para ser utilizado permanentemente por los servicios de prevención del suicidio.

Quizás las funcionalidades con el clasificador que da mejor resultado detecten demasiados Falso Positivo (sobrecarga de los servicios) o detecten algunos casos de Verdadero Positivo, cuando lo que necesita el servicio de prevención son mensajes de personas que necesitan ayuda (y estos mensajes no llegan). Por estas razones, se decidió crear una medida bidimensional llamada Índice bidimensional de precisión (PBI) representado en el gráfico de calor en la *ilustración 10*, donde el eje Y representa la proporción de verdaderos positivos con respecto al total de todas las muestras clasificadas como “Riesgo” (Ecuación 7) y en el eje X la misma proporción, pero con la clase “Sin riesgo” (Ecuación 8).

De esta forma, los puntos que más interesan son los que se encuentran en la esquina superior derecha. Un punto ubicado en el borde de esta esquina representa la clasificación perfecta, donde tanto $\text{precisión}_{\text{riesgo}}$, como $\text{precisión}_{\text{no_risk}}$ para mensajes de riesgo serían óptimos. Sin embargo, como se ve en la *ilustración 10*, los resultados se acumulan principalmente en la parte inferior derecha (aquellos clasificadores que siempre se clasifican como “Sin riesgo”) y suben hasta un punto centrado en 0.8.

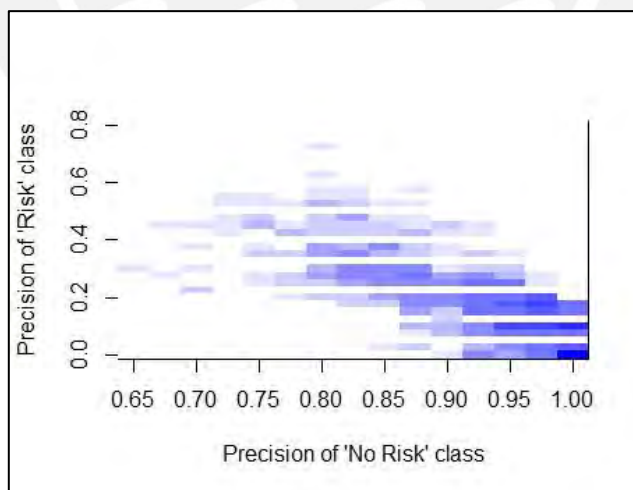


Ilustración 10 Representación del Índice Bidimensional de Precisión (PBI) de los resultados del Corpus con dos clases.

Este gráfico refleja, en el eje X, la clasificación de los mensajes de no riesgo, que no son de interés para los servicios de prevención y, en el eje Y, los mensajes con riesgo de ideación suicida o depresión, que serían los mensajes para enviar a dichos servicios. De esta forma, los puntos más interesantes para estos servicios serán los que se ubican

en la parte superior del gráfico y que están más a la derecha porque esto devolverá más mensajes de riesgo y pocos mensajes que no indiquen ningún peligro.

Por lo tanto, los servicios de prevención siempre elegirán el sistema anterior debido a la misma cantidad de mensajes falsos positivos, por lo que tendrán más positivos verdaderos. Pero dependerá de la carga de trabajo de los servicios de prevención que elegirán la columna porque si tienen muchas cargas de trabajo, estarán interesados en aplicar el sistema que menos errores cometa en el "Sin Riesgo" aunque algunos de los de "Riesgo" se escape. Los puntos ubicados en la parte superior se enumeran en la tabla 37. En todos estos sistemas, la característica de POS es relevante, siendo la única que aparece en todos.

Además, los árboles de decisión o variantes suelen estar en estos sistemas. El punto en el que el sistema obtiene la mayor precisión para la clase de riesgo es el que se encuentra en una precisión sin riesgo de 0,799, pero también es uno de los sistemas de la lista que utiliza más funciones. Esto demuestra que, si bien el POS es la característica que mejor discrimina los tipos de mensajes, agrega características que ayudan a la precisión de clasificación de la clase "Riesgo", aunque la categoría "Sin Riesgo" empeora.

Características del Corpus	Clasificador	Precisión No Riesgo	Precisión Riesgo
POS,SYNSETS,lemma,word	RandomCommitte	0,958	0,310
POS,stem	RandomTree	0,944	0,345
stopwords,POS,stem,SYNSETS,word,NUM	RandomTree	0,930	0,414
stopwords,POS,stem,SYNSETS,lemma,word,NUM	RandomizableFiltered	0,915	0,448
POS,SYNSETS,NUM	KStar	0,901	0,448
POS,stem,NUM	RandomTree	0,901	0,414
POS,stem,SYNSETS,NUM	RandomTree	0,887	0,517
stopwords,POS,stem	RandomizableFiltered	0,887	0,448
POS,NUM	RandomTree	0,873	0,586
stopwords,stem	RandomizableFiltered	0,845	0,517
POS,SYNSETS,word,NUM	RandomizableFiltered	0,831	0,517
stem,lemma,NUM	RandomizableFiltered	0,831	0,586
stopwords,POS,SYNSETS	RandomizableFiltered	0,817	0,517
stopwords,POS,stem,NUM	RandomizableFiltered	0,803	0,621
stopwords,POS,stem,word,lemma,word,NUM	RandomizableFiltered	0,789	0,724

Tabla 37 Los clasificadores y características utilizados para los modelos de aprendizaje más relevantes siguen el PBI.

7.1.4. Discusión

Los algoritmos de aprendizaje automático que han obtenido mejores resultados son KStar y RandomCommittee cuando se han utilizado con funcionalidades con mayor capacidad de generalización, como el Part-Of-Speech (POS), el identificador semántico del término (SYNSET) y la sustitución de números por un identificador único (NUM). Las

características como la raíz de las palabras (STEM), el lema (LEMMA) y la bolsa de palabras (WORD) han sido relegadas a posiciones inferiores, probablemente porque el Corpus aún es demasiado pequeño para que estas características aparezcan tanto en los textos de entrenamiento como en los textos de prueba.

Cabe señalar que la eliminación de palabras comunes (stopwords) como preposiciones, determinantes, artículos, ciertos verbos, etc., tiende a dar peores resultados, probablemente porque algunos de los términos excluidos (como 'no', 'a lot', 'be', ...), en combinación con otras características, son relevantes para evaluar el riesgo de ideación suicida y, por tanto, no es aconsejable descartarlas.

Por otro lado, durante este trabajo se ha determinado que las medidas habituales para evaluar la calidad de los resultados en sistemas desequilibrados (F-measure, ROC Área) no son las más adecuadas para la prevención del suicidio. Ya que en estas circunstancias el clasificador que proporciona el mejor resultado puede detectar, por desequilibrio, muchos falsos positivos (sobrecarga de los servicios) o detectar algunos casos de verdaderos positivos. Cuando lo que necesita el servicio de prevención son mensajes de personas que necesitan ayuda, según las necesidades de cada centro de prevención o según el volumen de mensajes.

Por eso se propuso una nueva medida denominada ***Precision and Recall Bidimensional Index*** (PBI), que nos muestra en un gráfico bidimensional (Ilustración 10), ***la precisión de clasificar un mensaje como Riesgo con respecto a la cobertura de mensajes de riesgo***, de tal forma que los sistemas que aparecen en la parte más alta y más a la derecha de la figura, serán, en general, los más interesantes, pudiendo aumentar el número de mensajes con riesgo de ideación suicida, que reciben los centros a costa de recibir más mensajes que no tienen relación con el suicidio o viceversa, dependiendo de la carga de trabajo en un momento dado de estos servicios.

Esto implica que elegir el sistema con el valor más alto de F-measure o ROC Área, no siempre será el más adecuado. Lo que está claro es que siempre se deben elegir aquellos sistemas que se encuentran en la parte superior de la figura, siendo peores sistemas los que están por debajo de otro sistema en la misma columna.

7.2. Experimentación 2

El objetivo de esta investigación fue determinar un método semi-supervisado para poblar el corpus Life, utilizando una técnica de bootstrapping. Con ello, hemos intentado mejorar la detección y clasificación automática de textos extraídos de redes sociales y foros relacionados con el suicidio y la depresión en base al corpus Life [208].

7.2.1. Anotación semi-supervisada de mensajes potencialmente suicidas en un corpus supervisado

En trabajos previos con el Corpus Life [112][173], los autores utilizaron técnicas de aprendizaje automático para analizar sistemáticamente todas las combinaciones posibles de características textuales. Probaron 28 algoritmos clasificadores supervisados usando diferentes características del corpus. El estudio concluyó que sería interesante aumentar el corpus para mejorar el rendimiento.

El Corpus Life constaba originalmente de 102 mensajes suicidas, (71 textos en inglés y 31 textos en español) 70 muestras (Sin riesgo) y 32 textos (Riesgo), divididos en 4 clases Sin Riesgo, Urgente, Posible e Inmediato, todos ellos desequilibrados. Los resultados obtenidos en los experimentos realizados no fueron los mejores (Tabla 38).

Alert Level	Quantity	EN	ES
No risk	70 (68,6%)	45 (63,4%)	25 (80,6%)
Urgent	19 (18,6%)	15 (21,1%)	4 (12,9%)
Possible	8 (7,8%)	6 (8,5%)	2 (6,5%)
Immediate	5 (4,9%)	5 (7%)	0 (0%)

Tabla 38. Número de muestras para cada tipo de "Nivel de alerta".

Como el corpus era muy pequeño y había demasiadas categorías para obtener datos estadísticamente significativos [112], se decidió fusionar las tres clases de riesgo (Posible, Urgente e Inmediato) en una sola, manteniendo intacta la clase Sin riesgo, para reducir el desequilibrio. y por tanto la calidad de los experimentos. En este trabajo tomamos la misma decisión manteniendo las dos categorías Risk y No Risk, que aparecen en la Tabla 39.

Alert Level	Quantity	EN	ES
No risk	70 (68,63%)	45	25
Risk	32 (31,37%)	26	6

Tabla 39. Número de muestras para cada tipo de "Nivel de alerta".

Para incrementar el número de muestras en el corpus, decidimos recolectar textos de la plataforma social Reedit [223], dentro de la cual se encuentra el subreddit "SuicideWatch", que a la fecha de extracción contaba con 984 textos en idioma inglés, con el fin de aumentar el corpus y observar su funcionamiento. Estos textos fueron extraídos usando la biblioteca PRAW (acrónimo de "Python Reddit API Wrapper") en Python, que permite el acceso a Reddit a través de una cuenta de desarrollador [93]. Una vez extraídos los 984 textos, se preprocesaron, y se realizaron experimentos de acuerdo con el esquema que se puede visualizar en la Ilustración 11.

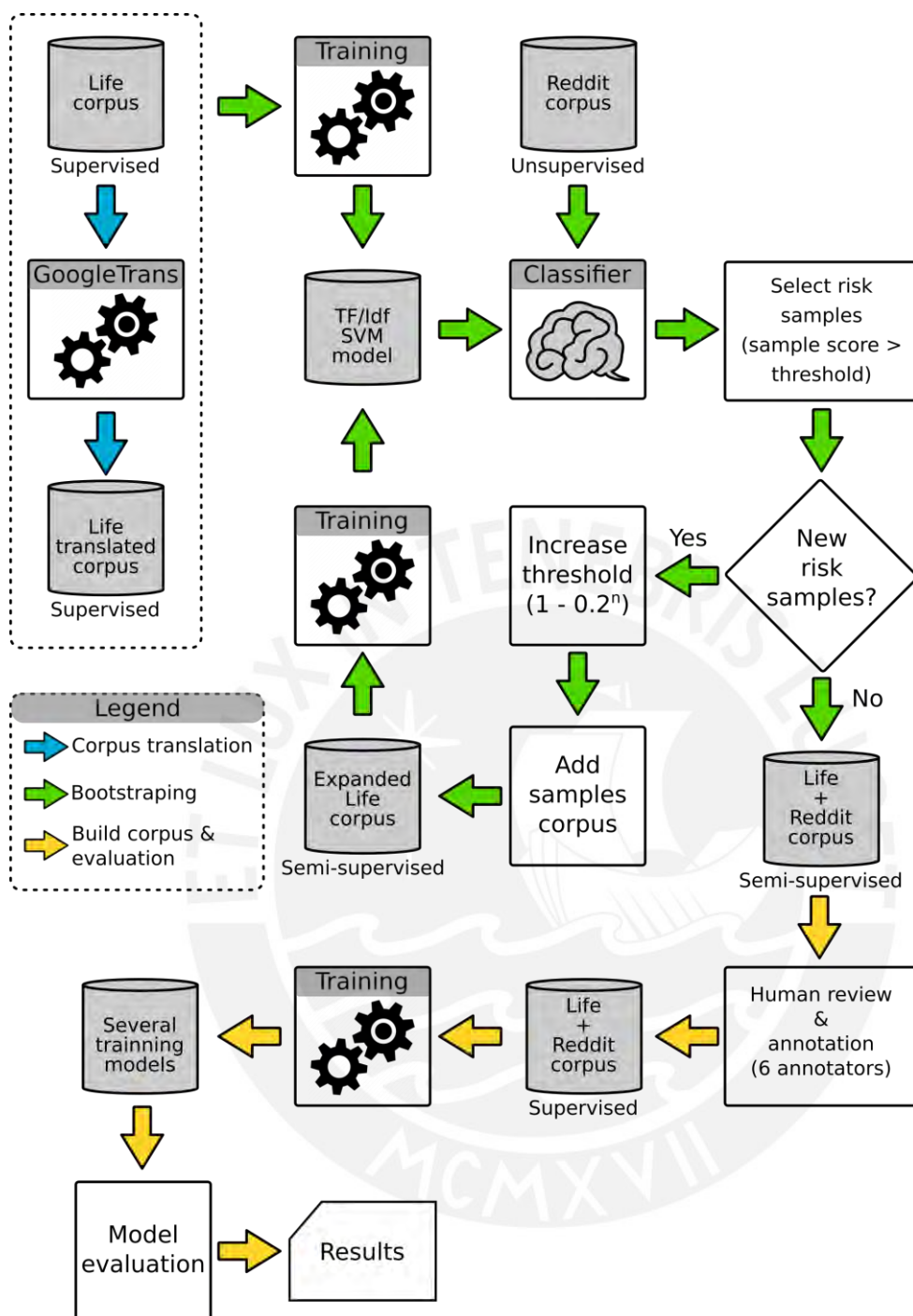


Ilustración 11. Esquema de flujo de trabajo del sistema. El sistema se evaluó utilizando el Life Corpus original y el Corpus Life traducido. El sistema se compone de tres procesos: i) proceso de traducción, ii) arranque de la expansión del corpus, iii) revisión, construcción y evaluación del corps supervisado.

Con el Corpus Life supervisado original y el nuevo Reddit Corpus no supervisado, desarrollamos el sistema que se muestra en la Ilustración 11, para incrementar las muestras del Corpus Life, especialmente aquellas etiquetadas como de Riesgo.

Luego de esto, utilizando las muestras del Corpus Life (inglés + español traducido al inglés), se creó un clasificador basado en el algoritmo SVM y utilizando las características de BoW con ponderación de términos Tf/Idf [224]. Este clasificador inicial se utilizó para aumentar el corpus con una técnica de Bootstrapping Uncertainty

Sampling. Como el Corpus Life estaba muy desequilibrado hacia muestras sin riesgo (Tabla 39), sólo nos interesaba elegir las muestras que el clasificador categorizaba como Riesgo.

De esta forma, considerando que el clasificador tendría una tasa de error importante, también se incluirían muestras sin riesgo. En cada iteración, el umbral de corte se incrementó exponencialmente para limitar la aceptación de nuevas muestras comparando ese umbral con la puntuación de confianza otorgada por el clasificador sklearn SVM para cada muestra. Eso significa que, si tengo la muestra X y Y, y el puntaje de SVM para la muestra X es más alto que el umbral, y el puntaje de la muestra Y es más bajo, se agregó la muestra X y se rechazó la muestra Y en esa iteración. Específicamente, el umbral creció en cada iteración $(1-0.2^n)$, donde n es el número de iteración. Repetimos esta iteración hasta que cualquiera de los puntajes de confianza de SVM para las muestras evaluadas superó el umbral.

En cada iteración del bootstrapping, las muestras aceptadas no supervisadas se añadían al corpus supervisado, entrenando de nuevo el modelo con las nuevas muestras y repitiendo la iteración hasta que, como se ha mencionado, el clasificador dejaba de clasificar muestras no etiquetadas previamente como en riesgo. La Tabla 40 muestra el número de muestras de corpus de cada iteración y el nuevo umbral calculado.

Interacción	Numero de ejemplos	Limite`
1	102	0,8
2	225	0,96
3	302	0,992

Tabla 40. Numero de Ejemplos y limites por iteración

El proceso de arranque anterior nos dejó con un corpus de 302 muestras de Reddit, 200 de estas muestras fueron evaluadas por seis anotadores divididos en cuatro grupos de 50 muestras, de modo que uno de los anotadores evaluó las muestras de los cuatro grupos, mientras que cinco anotadores independientes anotaron sólo su grupo de 50 muestras. Se alcanzó el acuerdo mutuo de 0,86, lo que representa 171 acuerdos en la evaluación de los textos [121]. Los resultados de la anotación se muestran en la Tabla 41.

Group	Reviewer	Suicide text Risk / No Risk	Mutual-agreement TP / TN	Kappa Cohen
1	Reviewer GC	38 / 12	38 / 3	0,82
	Reviewer RA	47 / 3		
2	Reviewer CC	46 / 4	46 / 4	1,00
	Reviewer RA	46 / 4		
3	Reviewer KM	45 / 5	41 / 3	0,88
	Reviewer RA	47 / 3		
4	Reviewer AR	32 / 18	26 / 10	0,72
	Reviewer JG	31 / 19		
	Reviewer RA	35 / 15		
Totales		364 / 87	151 / 20	0,86

Tabla 41. Acuerdos entre revisores

Para asegurar la calidad de los datos, de estas 302 muestras se mantuvieron las 171 en las que todos los anotadores coincidieron, es decir, hubo mutuo acuerdo en la clasificación de dichas muestras. Estas muestras se unieron a las muestras del corpus de Life para construir cinco corpus diferentes de diferentes tamaños (Tabla 42).

Corpus	Risk	Not risk	Total
Corpus in English language			
Life	30	72	102
Reddit	153	18	171
Life+Reddit	183	90	273
Corpus in Spanish English			
Life_es_en	30	72	102
Life_es_en+Reddit	183	90	273

Tabla 42. Corpus utilizado en experimentos.

Los experimentos del presente trabajo se realizaron con estos datos. Las razones por las que usamos 5 corpus fueron las siguientes: i) las fuentes de datos eran muy heterogéneas y tuvimos que verificar que los resultados no se deterioraran usando el corpus separado; y ii) como podemos ver en la Tabla 39, el Corpus Life contiene una mezcla de mensajes en inglés y español y queríamos probar cómo la traducción automática afectaba el desempeño del sistema con un corpus donde todos los mensajes estaban en el mismo idioma con traducción, o conservando los mensajes originales. No hay opción del corpus de Reddit con y sin traducción porque, en el corpus de Reddit,

todos los mensajes ya estaban en inglés. Por tanto, repetimos todos los experimentos con:

Life: solo la versión traducida de Life.

Reddit: solo el corpus de Reddit que como muestra contenía los textos resultantes del acuerdo mutuo entre los anotadores.

Life+Reddit: la combinación de muestras del corpus Life traducidas al inglés y las muestras de Reddit de mutuo acuerdo.

Life_es_en: El corpus de Life original y sin traducir, con muestras mixtas en inglés y español.

Life_es_en+Reddit: el corpus de Life original sin traducir más muestras con el acuerdo mutuo de Reddit.

Con estas combinaciones, pudimos evaluar la mejora del rendimiento cuando agregamos el corpus de Reddit, o cómo la traducción automática de los textos afectó este rendimiento. La traducción se realizó a través de Google Translation [133].

Una vez que se obtuvieron los diferentes corpus, usamos dos clasificadores diferentes: Support Vector Machine (SVM) y clasificador de intención RASA. El algoritmo de clasificación SVM se utiliza ampliamente en la clasificación de textos, dando buenos resultados en diferentes procesos de investigación [225]. Para esta máquina de aprendizaje se utilizó su implementación por medio de la biblioteca de aprendizaje automático sklearn [226]. Además, también queríamos realizar experimentos con algoritmos de aprendizaje profundo, pero debido al pequeño tamaño del corpus, decidimos utilizar el algoritmo de comprensión del lenguaje natural RASA (NLU), que permite clasificaciones utilizando modelos de lenguaje junto con técnicas de aprendizaje profundo [227].

Con el algoritmo RASA utilizamos las características que están definidas por defecto en el algoritmo Lexical Syntactic Featurizer: "low", "title", "upper," BOS "," EOS "," digit "," pos ". Low indica si el término está en minúscula o no, title si la palabra es un título, upper si la palabra está en mayúscula, digit si es un número, pos es parte del discurso, BOS si es una sentencia y EOS si es el fin de la sentencia.

Por otro lado, para los algoritmos de SVM se utilizaron las siguientes características: Bolsa de Palabras (BoW) con y sin Término-Frecuencia / Frecuencia de Documento Inverso (Tf / Idf), como ponderador de términos, y con o sin palabras vacías. Para tratar de mejorar la cobertura de los resultados, utilizamos incrustaciones de palabras para expandir los términos de cada mensaje. Para el uso de incrustaciones de palabras usamos la biblioteca Polyglot [228], que, dado un término, sugiere un número n de términos cercanos a dicho término cuyo vector de incrustación es menor que una

distancia d del vector del término de búsqueda. Después de varias pruebas preliminares, decidimos establecer n en un valor de 10 y d en un valor de 0,85.

Evaluamos los resultados con las siguientes métricas: precisión simple, precisión equilibrada, micro f1, macro f1, f1 ponderada, micro precisión, precisión macro, precisión ponderada, micro recuperación, recuperación macro, recuperación ponderada, micro jaccard, macro jaccard, jaccard ponderado, como el corpus estaba desequilibrado, ya sea hacia muestras sin riesgo en el caso del Life Corpus o muestras de riesgo para el Reddit o Life+Reddit Corpus, decidimos utilizar macro f1 como nuestra métrica principal, ya que es la que mejor respondió a el corpus desequilibrado, que calcula el estadístico f1, separado por clases y no utiliza pesos para agrupar [229].

Los experimentos se realizaron utilizando la técnica de validación cruzada de 10 veces y se repitieron 30 veces, con diferentes divisiones aleatorias de validación cruzada, para obtener significación estadística utilizando una prueba t para diferencias de medias de los valores de macro f1.

Accuracy. - Puntuación de clasificación de precisión, calcula la precisión de subconjunto [230].

Simple accuracy. - El conjunto de etiquetas predichas para un ejemplo debe coincidir exactamente con el conjunto correspondiente de etiquetas en y_true .

$$accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(\hat{y}_i = y_i)$$

Ecuación 14. Accuracy

Balanced accuracy. - Precisión equilibrada, se utiliza en problemas de clasificación binaria y multiclase para tratar con conjuntos de datos desequilibrados. Se define como el promedio de recuperación obtenido en cada clase. El mejor valor es 1 y el peor valor es 0 cuando $adjuste=False$.

$$balanced_accuracy = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right)$$

Ecuación 15. Balanced_accuracy

F1 score. - Conocida también como F-score o F-measure, es un promedio ponderado de precisión y recall, donde una puntuación de F1 alcanza su mejor valor en 1 y peor puntuación en 0. La contribución relativa de precisión y recall a la puntuación F1 es igual, intuitivamente la precisión es la capacidad del clasificador de no etiquetar como positiva una muestra que es negativa, y recall representa la capacidad del clasificador para encontrar todas las muestras positivas [230].

Micro_F1. - Calcula las métricas globalmente contando el total de positivos verdaderos, falsos negativos y falsos positivos.

$$\text{Micro } F_1\text{-score} = 2 \frac{\text{Micro-precision} * \text{Micro-recall}}{\text{Micro-precision} + \text{Micro-recall}}$$

Ecuación 16. Micro F1-score

Macro_F1.- Calcula las métricas para cada etiqueta y obtiene la media no ponderada. Esta medida no tiene en cuenta el desequilibrio de etiquetas.

$$\text{Macro } F_1\text{-score} = \frac{1}{N} \sum_{i=0}^N F1\text{-score}_i$$

Ecuación 17. Macro F1-score

Weighted_F1.- Calcula las métricas para cada etiqueta y obtiene un promedio ponderado por soporte (Numero de instancias verdaderas para cada etiqueta). Esto altera al macro para tener en cuenta el desequilibrio de etiquetas y puede dar lugar a una puntuación F que no esté entre precision y recall.

Precision. - La precisión mide cuantas veces el sistema ha identificado un tema correctamente, entre todas las veces que ese tema ha sido identificado por el sistema. Se puede considerar a la precisión como una puntuación derivada de la capacidad de la máquina para extraer temas de un texto dado y su capacidad para asegurarse de que muchos de estos temas extraídos sean correctos [230].

Micro_precision. - Calcula métricas globales contando el total de verdaderos positivos, falsos negativos y falsos positivos.

$$\text{Micro - precision} = \frac{TP1 + TP2}{TP + FP1 + TP2 + FP2}$$

Ecuación 18. Micro_precision

Macro_precision. - Calcula las métricas para cada etiqueta y obtiene la media no ponderada. Esta medida no tiene en cuenta el desequilibrio de etiquetas.

$$\text{Macro - precision} = \frac{\text{Precision1} + \text{Precision2}}{2}$$

Ecuación 19. Macro_precision

Weighted_precision. - Calcula las métricas para cada etiqueta y obtiene un promedio ponderado por soporte (Numero de instancias verdaderas para cada etiqueta). Esto altera macro para tener en cuenta el desequilibrio de etiquetas y puede dar lugar a una puntuación F que no esté entre precision y recall.

Recall. - En palabras simples la recuperación en el análisis de textos mide cuantas veces el sistema es capaz de identificar un tema específico de un texto no estructurado, sobre la cantidad total de veces que el tema realmente se menciona en ese texto. Es la relación entre el número de verdaderos positivos y el número de falsos negativos. Recall es intuitivamente la capacidad del clasificador para encontrar todas las muestras positivas [230].

Micro_recall. - Calcula métricas globales contando el total de verdaderos positivos, falsos negativos y falsos positivos.

$$\text{Micro - recall} = \frac{TP1 + TP2}{TP + FP1 + TP2 + FP2}$$

Ecuación 20. Micro_recall

Macro_recall. - Calcula las métricas para cada etiqueta y obtiene la media no ponderada. Esta medida no tiene en cuenta el desequilibrio de etiquetas.

$$\text{Macro - Recall} = \frac{\text{Recal1} + \text{Recall2}}{2}$$

Ecuación 21. Macro_recall

Weighted_recall. - Calcula las métricas para cada etiqueta y obtiene un promedio ponderado por soporte (Numero de instancias verdaderas para cada etiqueta). Esto altera macro para tener en cuenta el desequilibrio de etiquetas y puede dar lugar a una puntuación F que no esté entre precision y recall.

Jaccard. - El índice coeficiente Jaccard, o coeficiente de similitud Jaccard, definido como el tamaño de la intersección dividido por el tamaño de la unión de dos conjuntos de etiquetas, se utiliza para para comparar el conjunto de etiquetas predichas para una muestra con el conjunto correspondiente de etiquetas en y true [230].

$$J(y, \hat{y}_i) = \frac{|y_i \cap \hat{y}_i|}{|y_i \cup \hat{y}_i|}$$

Ecuación 22. Jaccard similarity

Micro_jaccard. - Calcula métricas globales contando el total de verdaderos positivos, falsos negativos y falsos positivos.

Macro_jaccard. - Calcula las métricas para cada etiqueta y obtiene la media no ponderada. Esta medida no tiene en cuenta el desequilibrio de etiquetas.

Weighted_jaccard. - Calcula las métricas para cada etiqueta y obtiene un promedio ponderado por soporte (Numero de instancias verdaderas para cada etiqueta). Esto altera macro para tener en cuenta el desequilibrio de etiquetas y puede dar lugar a una puntuación F que no esté entre precision y recall.

Como el corpus estaba desequilibrado, ya sea hacia muestras sin riesgo en el caso del corpus Life, o en muestras de riesgo para el corpus Reddit o Life+Reddit, decidimos seleccionar a **Macro F1** como nuestra métrica principal, que es la que mejor respondió al corpus desequilibrado, calcula el estadístico F1, separado por clases y no utiliza ponderaciones para la agrupación [229].

Los experimentos se realizaron utilizando la técnica de validación cruzada de 10-folds y se repitieron 30 veces, con diferentes divisiones de validación cruzada, para obtener significación estadística.

7.2.2. Resultados

Como teníamos dos variantes del Corpus Life, una con todas las muestras traducidas al inglés y la otra manteniendo las muestras en español, decidimos dividir los experimentos con todo el corpus que tenía una variante y con los que no. Por ello, para la primera parte de esta sección realizamos los experimentos con las colecciones Life, Reddit y Life + Reddit con todos los textos traducidos al inglés. Sin embargo, en la segunda parte, usamos el corpus Life_es_en y Life_es_en + Reddit, con el corpus Life sin traducir. Para la segunda parte no usamos el corpus de Reddit porque está completamente en idioma inglés.

7.2.2.1. Experimentos corpus en idioma Ingles

En la Ilustración 12, podemos observar los resultados usando el clasificador SVM (usando las características BoW, BoW + Embeddings, Tf / Idf y Tf / Idf + Embeddings) y Rasa (con el sistema de extracción de características predeterminado) para las tres diferentes colecciones de datos en inglés: Life, Reddit y Life + Reddit.

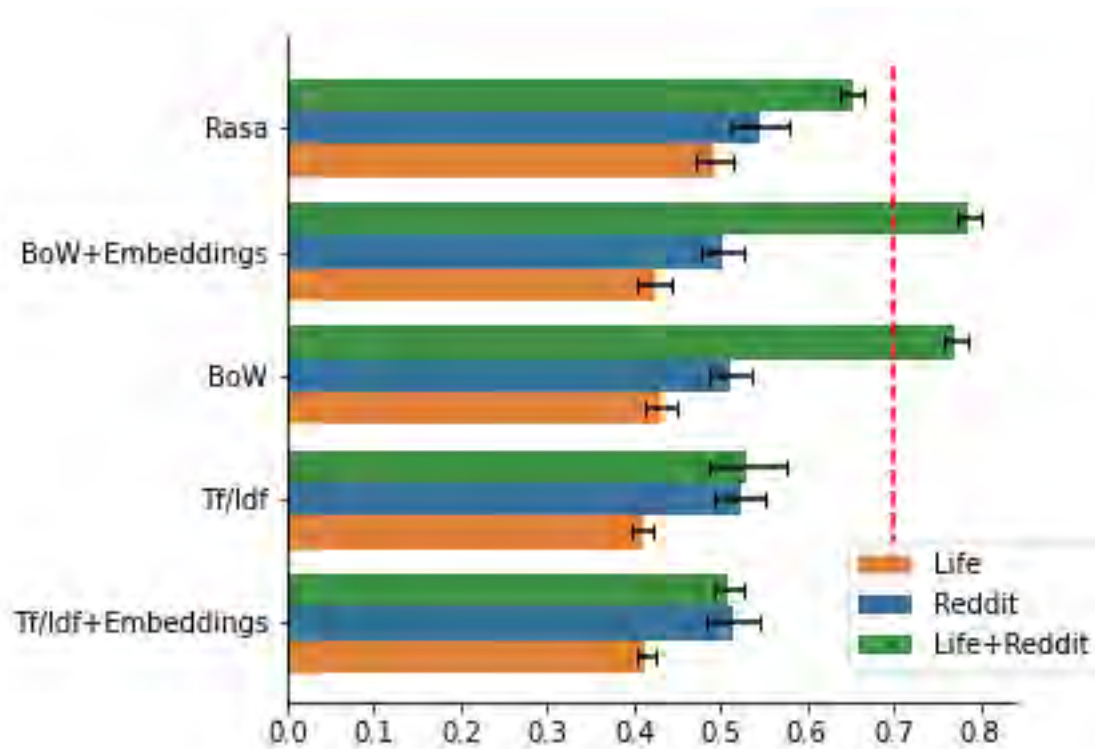


Ilustración 12. La línea puntiaguda vertical es el resultado original de la medida f con el Corpus Life

Como se puede observar en la Ilustración 12, los resultados de macro F1 fueron mejores con el clasificador SVM y BoW como características (sin el pesador Tf / Idf). No hubo una diferencia estadísticamente significativa si usamos BoW con o sin incrustaciones de palabras. Por tanto, ambos sistemas parecen tener un rendimiento similar. Sin embargo, para corpus más pequeños, el sistema Rasa fue el mejor sistema, confirmando que este clasificador alcanza buenos resultados, cuando el número de muestras es pequeño, debido a que incluye modelos de aprendizaje.

En la Tabla 43 podemos ver los resultados con más detalle y observar a las medidas **Macro Precision** alcanzar un valor de 0,80, mientras que **Macro F1** y **Macro Recall**, alcanzan un valor de 0,79. Los resultados en color representan cuando los valores son estadísticamente significativos ($P < 0,01$) con respecto al mejor valor de cada corpus evaluado. Como podemos ver en la Tabla 42, el mejor sistema para cada corpus tiene mejor desempeño en las tres medidas (Macro F1, Macro Precisión y Macro Recall).

Features	Macro F1	Macro Precision	Macro Recall
Life			
Rasa	0,49±0,02	0,52±0,03	0,53±0,02
BoW	0,43±0,02	0,40±.03	0,50±0,01
BoW+Embedding	0,42±0,02	0,39±.03	0,51±0,02
Tf/Idf	0,41±0,01	0,35±.01	0,51±0,01

Tf/Idf+Embeddings	0,41±0,01	0,35±0,01	0,51±0,01
Reddit			
Rasa	0,55±0,03	0,54±0,04	0,57±0,03
BoW	0,51±0,02	0,48±0,03	0,55±0,02
BoW+Embedding	0,50±0,02	0,47±0,02	0,54±0,02
Tf/Idf	0,52±0,03	0,50±0,03	0,55±0,03
Tf/Idf+Embeddings	0,52±0,03	0,49±0,03	0,55±0,03
Life+Reddit			
Rasa	0,65±0,01	0,76±0,02	0,66±0,01
BoW	0,77±0,01	0,77±0,01	0,79±0,01
BoW+Embeddings	0,79±0,01	0,80±0,01	0,79±0,01
Tf/Idf	0,53±0,04	0,72±0,06	0,58±0,05
Tf/Idf+Embeddings	0,51±0,02	0,68±0,04	0,56±0,01

Tabla 43. Macro F1, Macro Precision, y Macro Recall. Corpus en idioma inglés combinado con características de entrenamiento. El intervalo de confianza se calculó con $p < 0,01$

7.2.2.2. Experimentos corpus en idiomas español e inglés

Como se mencionó anteriormente, en el segundo grupo de experimentos, el corpus original de Life se usó sin traducir ninguna muestra o en combinación con Reddit, que está completamente en inglés. El objetivo de estos experimentos fue observar si la traducción afectaba o no a los resultados en la clasificación de mensajes con ideación suicida. Como en experimentos anteriores, se utilizaron los clasificadores SVM y Rasa. Para el primero, las características de BoW se extrajeron de cada corpus, BoW se expandió con incrustaciones de palabras y sus variantes de peso con Tf/Idf. El resultado de estos experimentos se puede ver en la Ilustración 13.

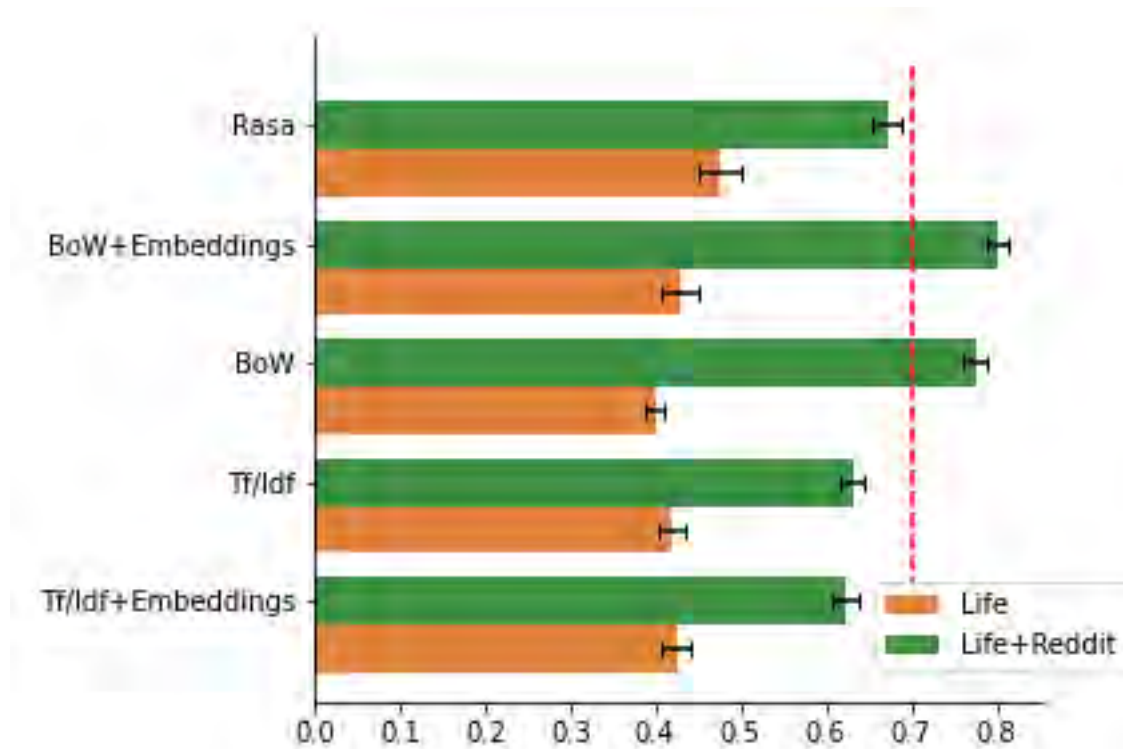


Ilustración 13. La línea puntiaguda vertical es el resultado original de la medida f con el Corpus Life

Una vez más, los mejores resultados se obtuvieron utilizando el corpus más extenso (Life_es_en+Reddit) con el clasificador SVM y extrayendo las características de BoW, con o sin expansión de incrustación de palabras. Este sistema da una macro f1 de 0,80 ($p < 0,01$). El mejor resultado para el Life Corpus sin traducción fue usando Rasa (0,48). En la Tabla 44, presentamos el mejor sistema para cada corpus con mejor desempeño en las tres medidas (Macro F1, Macro Precisión y Macro Recall).

Features	Macro F1	Macro Precision	Macro Recall
Life_es_en			
Rasa	0,48±0,03	0,51±0,03	0,50±0,02
BoW	0,40±0,01	0,35±0,02	0,49±0,01
BoW+Embeddings	0,43±0,02	0,40±0,04	0,51±0,01
Tf/Idf	0,42±0,02	0,37±0,02	0,51±0,01
Tf/Idf+Embeddings	0,43±0,02	0,37±0,02	0,52±0,01
Life_es_en+Reddit			
Rasa	0,67±0,02	0,78±0,02	0,67±0,01
BoW	0,78±0,01	0,78±0,01	0,79±0,01
BoW+Embeddings	0,80±0,01	0,81±0,01	0,81±0,01

Tf/Idf	0,63±0,01	0,78±0,02	0,64±0,01
Tf/Idf+Embeddings	0,62±0,01	0,76±0,02	0,63±0,01

Tabla 44. Macro F1, Macro Precisión y Macro Recall. Corpus en idioma inglés y español combinado con funciones de formación. El intervalo de confianza se calculó con $p < 0,01$

Aunque la mayor precisión de corpus y la recuperación coinciden con los mejores resultados, esto no sucede con el corpus más pequeño. El mejor sistema de recuperación difiere del mejor sistema macro f1 ($p < 0,01$). Además, si comparamos estos resultados con el corpus traducido (tabla 42), podemos observar una ligera diferencia no significativa ($0,80 \pm 0,01$ vs $0,79 \pm 0,01$ $p < 0,01$). Esto quiere decir que el proceso de traducción automática con la librería GoogleTrans de las muestras en español del corpus original al idioma inglés no empeora significativamente el rendimiento de la detección de mensajes suicidas.

7.2.3. Discusión

Si bien O'Dea y Gómez desarrollaron corpus cuya calidad se evaluó a través de acuerdos, su metodología de desarrollo fue diferente: el primero, desarrolló un corpus a partir de datos descargados de Twitter en un mismo período, de los cuales el 14% seleccionado se dividió aleatoriamente en dos datos conjuntos para ser evaluados por codificadores humanos que los clasificaron en tres categorías: "Muy preocupante" (14%), "Posiblemente preocupante" (56%) y "Seguro de ignorar" (29%), con un acuerdo Kappa de Cohen de 0,76, mientras que el clasificador identificó correctamente el 80% de los tuits en la categoría "Muy preocupantes". El segundo, inicialmente desarrolló un corpus supervisado con anotaciones de diferentes fuentes que, por ser evaluado con cuatro categorías, tuvo una concordancia Kappa de Cohen moderada de 0,52 (promedio $k = 0,55$), en el proceso de incrementar el tamaño del muestras del Corpus, se utilizó una metodología semi-supervisada con textos del subreddit "SuicideWatch", lo que permitió ampliar el número de muestras de 102 a 273 (183 riesgo y 90 Sin riesgo) con una concordancia entre anotadores Kappa de Cohen de 0,86.

Como hemos visto en la Sección 7.2.2, los resultados son prometedores, siendo un sistema de aprendizaje semi-supervisado capaz de lograr un Macro F1 de 0,78 – 0,81, cercano al acuerdo mutuo alcanzado por revisores humanos (Kappa de Cohen de 0,86). Estos resultados llegaron cuando el Life Corpus aumentó de tamaño al sumar las 171 muestras del Reddit Corpus, en las que los anotadores llegaron a un acuerdo mutuo. Estos resultados también nos muestran que la metodología semi-supervisada de Bootstrapping Uncertainty Sampling elegida para ampliar el Corpus Life con nuevas muestras es válida y útil para mejorar los resultados del sistema automático de detección de mensajes de depresión o ideación suicida.

El algoritmo Rasa NLU, funciona mejor que el algoritmo SVM, para Corpus más pequeños, tal vez porque usa algunos modelos de lenguaje entrenados previamente. Sin embargo, BoW y SVM funcionan mejor que el enfoque de aprendizaje profundo de Rasa. Expandir el texto usando incrustaciones de palabras no afecta significativamente los resultados ($p < 0,01$) ni la traducción automática de los textos en español del Life Corpus al idioma inglés usando la biblioteca GoogleTrans.

Por otro lado, antes de utilizar esta metodología asumimos que no habría tantos mensajes suicidas en el subgrupo de Reddit y que el clasificador inicial no podría encontrar tantos mensajes de este tipo. Por tanto, el corpus final será más equilibrado. Sin embargo, el clasificador inicial para las técnicas de arranque utilizando un clasificador SVM preliminar funcionó mejor de lo que esperábamos. Esto nos ha llevado a pasar de un Life Corpus donde la mayoría de las publicaciones no tenían riesgo de suicidio a otro Life_es_en+Reddit Corpus, donde la mayoría de las publicaciones, según los revisores, tenían indicios de ideación suicida o depresión. (Tabla 42).



Capítulo VIII

8. Conclusiones

Este capítulo contiene las conclusiones de los dos grupos de experimentos así, como las conclusiones generales de la tesis.

8.1. Conclusiones experimentación 1

La OMS reconoce al suicidio como una prioridad de salud pública mundial. Los estados miembros de la OMS se han comprometido a reducir las tasas nacionales de suicidio en un 10% antes del año 2020. Para ello acordaron elaborar y poner en práctica estrategias nacionales integrales de prevención del suicidio, para mejorar sus sistemas de información, actualizar sus datos científicos y realizar investigaciones colaborativas con universidades en salud mental.

El alto tráfico diario de información que se produce en redes sociales como Facebook (4 nuevos petabytes de datos) y Twitter (500 millones de Tweets) [166]. Es utilizado en procesos de investigación, en los que se han recopilado mensajes con ideas suicidas de redes sociales como Twitter, Weibo, Netlog y otros microblogs. Información esencial para tratar de disminuir la tasa de suicidio global anual estandarizada de 11,4 por 100.000 habitantes (911.375 suicidios en 2022).

Nos enfocamos en la evaluación de 28 algoritmos clasificadores supervisados con los parámetros predeterminados de Weka. Cada clasificador se evaluó mediante 124 combinaciones de características extraídas del Corpus Life. El clasificador que mostró un mejor rendimiento fue KStar con las funciones POS-SYNSET-NUM y POS-NUM. Suponemos que esto se debe al pequeño tamaño del Corpus porque estas características generalizan más que otras y pueden incluir en la clase correcta más muestras, pero con componentes léxicos o semánticos similares. Sin embargo, es menos probable que las características como una bolsa de palabras, raíces o lemas vean muestras con términos similares.

A pesar del pequeño tamaño del Corpus, hemos obtenido buenos resultados estadísticamente significativos, alcanzando 0,75 de medida F-measure y 0,81 de Área ROC con 2 clases y 0,72 de medida F-measure y 0,75 de Área ROC con 4 clases. Todos estos resultados son estadísticamente significativos con un margen de confianza de $p < 0,05$. Ambos resultados utilizan el Corpus Life con un micro acuerdo mutuo Kappa de Cohen moderado de 0,52. Aunque el Corpus tiene un acuerdo mutuo bajo, este se midió para comparar las 4 clases y no se midió para la separación de Riesgo y Sin Riesgo. Por tanto, se supone que este Corpus tendrá una mayor concordancia si se hubiera medido con solo dos clases. Sin embargo, los resultados son muy similares, y estos no varían mucho, ni con la diferencia de concordancia ni con los valores de la medida F-measure y del Área ROC.

Los resultados de esta evaluación y su análisis estadístico demuestran que el Corpus Life y algunas técnicas de aprendizaje automático podrían ser adecuados para detectar mensajes de ideación suicida a pesar de su pequeño tamaño. Para evaluar la confiabilidad de este Corpus, la concordancia observada, así como la estadística Kappa de Cohen se acompañaron de intervalos de confianza (IC) para proporcionar una

descripción de confiabilidad más detallada. Los IC proporcionan más información que indica un rango de valores (intervalo) que probablemente contenga el valor verdadero, con una probabilidad o un nivel de confianza. En nuestro caso, se han establecido intervalos de confianza de $p < 0,05$ y un nivel de significancia de $\text{Alpha} = 0,05$. Para verificar si nuestro acuerdo es estadísticamente significativo (es decir, $K \geq 0$), los intervalos de confianza proporcionados no deben incluir 0.

Otra ventaja de este trabajo, en comparación con otros, es el uso de un Corpus bilingüe (inglés y español), que no se enfoca en una sola red social, y está disponible gratuitamente bajo una licencia Creative Commons. Además, a diferencia de otros trabajos, el Corpus Life utilizado en esta investigación, está específicamente orientado a la detección de la ideación suicida y la depresión. Sin embargo, el Corpus también intenta detectar otras características indirectas, como emociones o sentimientos.

Es importante el hecho de que nuestros resultados proporcionan evidencia de que, dentro de los sistemas desequilibrados, las medidas habituales para evaluar la calidad de los resultados (medida F-measure, Área ROC) pueden no ser las más adecuadas para la prevención del suicidio, porque puede haber muchos positivos o verdaderos positivos, cuando lo que necesita el servicio de prevención son mensajes de personas que necesitan ayuda.

Hemos demostrado que, en general, una nueva medida denominada **Índice de precisión y recuperación bidimensional** (PIB), en el caso de sistemas desequilibrados, puede proporcionar mejores resultados que las medidas habituales para evaluar la calidad de los resultados (medida F-measure, área ROC), lo que aumenta el número de mensajes con riesgo de ideación suicida que reciben los centros a costa de recibir más mensajes que no tienen relación con el suicidio o viceversa.

El proceso de evaluación de todos estos sistemas supervisados ha sido bastante costoso tanto humana como computacionalmente. En esta investigación hemos utilizado los algoritmos de clasificación que vienen por defecto en Weka. En futuros trabajos, se evaluarán los textos con el estadístico Cohen's kappa previo a ser incluidos en el Corpus con solo dos características.

En esta investigación, usamos la fuerza bruta para ejecutar 3.472 experimentos diferentes, esto fue posible porque el Corpus era pequeño. Sin embargo, hay planes para expandir el Corpus considerablemente con nuevos datos provenientes de redes sociales y experimentar con nuevos algoritmos de clasificación y características. El estudio concluyó que es necesario incrementar el número de muestras del corpus para mejorar el rendimiento y alcanzar mejores resultados.

8.2. Conclusiones experimentación 2

Por haber sido anotados mediante una metodología supervisada la población del corpus de suicidios ha sido costosa. El objetivo de esta investigación fue incrementar el número de muestras del Corpus Life, utilizando un método semi-supervisado, que permite mantener la calidad de los textos agregados, reduciendo el esfuerzo humano.

Además, hemos demostrado dos cosas: i) la técnica Bootstrapping Uncertainty Sampling utilizada en el presente trabajo puede ser útil para incrementar el tamaño de un corpus adecuado para la prevención del suicidio utilizando enfoques de aprendizaje automático semi-supervisado, y ii) el Corpus Life ampliado puede ayudar a construir un

clasificador en el que se pueden detectar mensajes de ideación depresiva o suicida casi tan bien como el mutuo acuerdo humano.

Por medio de los experimentos realizados, hemos validado la metodología de anotación semi-supervisado que en esta investigación se utilizó para procesar muestras provenientes de redes de microblogging, blogs, foros u otras fuentes.

Se ha probado la metodología de traducción automática a través de la biblioteca GoogleTrans, para generar corpus paralelos expandidos a diferentes idiomas, desde los textos originales escritos principalmente en el idioma inglés, para realizar experimentos para la detección de usuarios suicidas en redes sociales en otros idiomas, realizadas por nuestro grupo de investigación o por otros grupos. Que podrán acceder al Corpus Life de forma gratuita bajo licencia Creative Commons en <https://github.com/PlataformaLifeUA>.

8.3. Conclusiones generales

El análisis de los resultados del primer grupo de experimentos concluyó que aumentar el número de muestras del corpus Life, era necesario para alcanzar mejores resultados.

Los resultados del segundo grupo de experimentos, muestran que la metodología semi-supervisado elegida, para incrementa el tamaño del Corpus Life, con nuevas muestras es válida y útil para mejorar el sistema automático de detección de mensajes suicidas. Siendo mejores los resultados con un Corpus con mayor cantidad de anotaciones.

Los resultados de los acuerdos de anotaciones manuales entre anotadores con el estadístico Cohen's kappa fue 0,86, y el mayor valor **Macro F1**, calculado de forma automática con la combinación **Life_es_en+Reddit +BoW_Embeddings** fue de 0,80, son altos, lo que permitirá con algoritmos de clasificación entrenar nuevos datos y realizar búsquedas de usuarios suicidas en redes sociales.

Por lo tanto, se verifica la hipótesis planteada en esta investigación.

H₁ Si se desarrolla un método semi-supervisado que detecte, clasifique y anote textos relacionados con el suicidio provenientes de entornos digitales entonces se incrementará el número de registros del Corpus de suicidio, lo que permitirá con algoritmos de clasificación entrenar nuevos datos y realizar búsquedas de usuarios suicidas en redes sociales.

Referencias Bibliográficas

- [1] Worldometer, "Worldometer - World statistics in real time." <https://www.worldometers.info/es/> (accessed Feb. 04, 2022).
- [2] "Digital 2022 - We Are Social Spain." <https://wearesocial.com/es/blog/2022/01/digital-2022/> (accessed May 20, 2022).
- [3] N. Rajkarnikar and D. Shrestha, "A Pilot Study on the Impact of ICT and New Media on Society and Culture," *SSRN Electronic Journal*, Nov. 2021, doi: 10.2139/SSRN.3997218.
- [4] T. Saito, J. Nakamura, and S. Yamazaki, "Development of Any from Middle English to Early Modern English: A Study Using the Helsinki Corpus of English Texts," in *English Corpus Linguistics in Japan*, Brill | Rodopi, 2016, pp. 211–223. doi: 10.1163/9789004334205_013.
- [5] Q. Liao, W. Wang, Y. Han, and Q. Zhang, "Analyzing the influential people in Sina Weibo dataset," in *GLOBECOM - IEEE Global Telecommunications Conference*, 2013, pp. 3066–3071. doi: 10.1109/GLOCOM.2013.6831542.
- [6] J. Weng, E. P. Lim, J. Jiang, and Q. He, "TwitterRank: Finding topic-sensitive influential twitterers," in *WSDM 2010 - Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*, 2010, pp. 261–270. doi: 10.1145/1718487.1718520.
- [7] M. A. Al-Garadi *et al.*, "Analysis of Online Social Network Connections for Identification of Influential Users," *ACM Comput Surv*, vol. 51, no. 1, pp. 1–37, Jan. 2018, doi: 10.1145/3155897.
- [8] IDC Corporate USA, "THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East," 2012. <https://www.idc.com/> (accessed Mar. 23, 2020).
- [9] C. Basu, H. Hirsh, and W. Cohen, "Recommendation as Classification: Using Social and Content-Based Information in Recommendation," 1998. Accessed: Oct. 13, 2020. [Online]. Available: <http://www.firefly.net>
- [10] M. Akpınar, B. Oral, D. Engin, ... E. E.-2019 4th, and undefined 2019, "A Semi-Automatic Annotation Interface for Named Entity and Relation Annotation on Document Images," *ieeexplore.ieee.org*, Accessed: Apr. 02, 2020. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/8907209/>
- [11] S. Raschka and V. Mirjalili, *Python machine learning : aprendizaje automático y aprendizaje profundo con Python, scikit-learn y TensorFlow*, Segunda. Marcombo, S.A., 2019.
- [12] A. J. Reagan, C. M. Danforth, B. Tivnan, J. R. Williams, and P. S. Dodds, "Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs," *EPJ Data Sci*, vol. 6, no. 1, p. 28, 2017, doi: 10.1140/epjds/s13688-017-0121-9.
- [13] X. Goldberg, "Introduction to semi-supervised learning," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 6, pp. 1–116, Jun. 2009, doi: 10.2200/S00196ED1V01Y200906AIM006.
- [14] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *ACM International Conference Proceeding Series*, 2006, vol. 148, pp. 161–168. doi: 10.1145/1143844.1143865.
- [15] T. Leung and F. L. Chung, "Persuasion driven influence propagation in social networks," in *ASONAM 2014 - Proceedings of the 2014 IEEE/ACM International Conference on*

- Advances in Social Networks Analysis and Mining*, Oct. 2014, pp. 548–554. doi: 10.1109/ASONAM.2014.6921640.
- [16] S. Choudhury and S. Pattnaik, “Emerging themes in e-learning: A review from the stakeholders’ perspective,” *Comput Educ*, vol. 144, p. 103657, Jan. 2020, doi: 10.1016/j.compedu.2019.103657.
- [17] A. Mircoli, C. Diamantini, and D. Potena, “Automatic Emotional Text Annotation Using Facial Expression Analysis,” 2017. Accessed: Apr. 01, 2020. [Online]. Available: <https://www.researchgate.net/publication/321579314>
- [18] A. O. Peschel, S. Kazemi, M. Liebichová, S. C. M. Sarraf, and J. Aschemann-Witzel, “Consumers’ associative networks of plant-based food product communications,” *Food Qual Prefer*, vol. 75, pp. 145–156, Jul. 2019, doi: 10.1016/j.foodqual.2019.02.015.
- [19] P. Rodriguez, A. Ortigosa, and R. M. Carro, “Extracting emotions from texts in E-learning Environments,” in *Proceedings - 2012 6th International Conference on Complex, Intelligent, and Software Intensive Systems, CISIS 2012*, 2012, pp. 887–892. doi: 10.1109/CISIS.2012.192.
- [20] I. Stefanova and A. Kiryantsev, “Analysis of User Groups in Social Networks to Detect Socially Dangerous People,” in *2018 International Scientific-Practical Conference on Problems of Infocommunications Science and Technology, PIC S and T 2018 - Proceedings*, Jan. 2019, pp. 789–792. doi: 10.1109/INFOCOMMST.2018.8632086.
- [21] S. R. Braithwaite, C. Giraud-Carrier, J. West, M. D. Barnes, and C. L. Hanson, “Validating Machine Learning Algorithms for Twitter Data Against Established Measures of Suicidality.,” *JMIR Ment Health*, vol. 3, no. 2, p. e21, May 2016, doi: 10.2196/mental.4822.
- [22] B. Desmet and V. Hoste, “Emotion detection in suicide notes,” *Expert Syst Appl*, vol. 40, no. 16, pp. 6351–6358, Nov. 2013, doi: 10.1016/J.ESWA.2013.05.050.
- [23] M. Van Egmond, R. D.-C. T. J. of Crisis, and undefined 1990, “The predictability of suicidal behavior: the results of a meta-analysis of published studies.,” *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, pp. 57–84., 1990.
- [24] D. D. Luxton, J. D. June, and J. M. Fairall, “Social media and suicide: a public health perspective.,” *Am J Public Health*, vol. 102 Suppl 2, no. SUPPL. 2, pp. S195-200, May 2012, doi: 10.2105/AJPH.2011.300608.
- [25] B. O’Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, “Detecting suicidality on Twitter,” *Internet Interv*, vol. 2, no. 2, pp. 183–188, May 2015, doi: 10.1016/j.invent.2015.03.005.
- [26] S. Velupillai, G. Hadlaczky, ... E. B.-G.-F. in, and undefined 2019, “Risk Assessment Tools and Data-Driven Approaches for Predicting and Preventing Suicidal Behavior,” *ncbi.nlm.nih.gov*.
- [27] L. Zhang, X. Huang, T. Liu, A. Li, Z. Chen, and T. Zhu, “Using linguistic features to estimate suicide probability of Chinese microblog users,” in *International Conference on Human Centered Computing*, 2014, pp. 549–559. doi: 10.1007/978-3-319-15554-8.
- [28] S. Abrutyn, A. S. Mueller, and M. Osborne, “Rekeying Cultural Scripts for Youth Suicide: How Social Networks Facilitate Suicide Diffusion and Suicide Clusters Following Exposure to Suicide,” *Soc Ment Health*, vol. 10, no. 2, pp. 112–135, Jul. 2020, doi: 10.1177/2156869319834063.
- [29] Y. Xiao and M. A. Lindsey, “Adolescent social networks matter for suicidal trajectories: disparities across race/ethnicity, sex, sexual identity, and socioeconomic status,” *Psychol Med*, pp. 1–12, 2021, doi: 10.1017/S0033291721000465.

- [30] A. Canossa, A. Azadvar, C. Hartevelde, A. Drachen, and S. Deterding, "Influencers in multiplayer online shooters evidence of social contagion in playtime and social play," in *Conference on Human Factors in Computing Systems - Proceedings*, May 2019, pp. 1–12. doi: 10.1145/3290605.3300489.
- [31] M. Scatà, A. Di Stefano, A. La Corte, and P. Liò, "Quantifying the propagation of distress and mental disorders in social networks," *Sci Rep*, vol. 8, no. 1, p. 5005, Dec. 2018, doi: 10.1038/s41598-018-23260-2.
- [32] G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield, "Analysing the connectivity and communication of suicidal users on twitter," *Comput Commun*, vol. 73, pp. 291–300, Jan. 2016, doi: 10.1016/J.COMCOM.2015.07.018.
- [33] N. Masuda, I. Kurahashi, and H. Onari, "Suicide Ideation of Individuals in Online Social Networks," *PLoS One*, vol. 8, no. 4, p. e62262, Apr. 2013, doi: 10.1371/journal.pone.0062262.
- [34] M. Hernandez and J. Pontes, "Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)," *the 5th Information Systems Research ...*, 2014.
- [35] Y. Huang, T. Goh, and C. Liew, "Hunting suicide notes in web 2.0-preliminary findings," *Multimedia Workshops, 2007.*, 2007.
- [36] P. Matykiewicz, W. Duch, and J. Pestian, "Clustering semantic spaces of suicide notes and newsgroups articles," *of the Workshop on Current Trends ...*, 2009.
- [37] J. Pestian and P. Matykiewicz, "Using natural language processing to classify suicide notes," *Proceedings of the*, 2008.
- [38] Organización Mundial de la Salud, "Plan de Acción sobre Salud Mental 2013 - 2020," 2013. doi: 978 924 250602 0.
- [39] World Health Organization, *Preventing suicide: A global imperative*. World Health Organization, 2014.
- [40] H. Jeong, B. Nam, S. J. Jo, W. C. Lee, and H. W. Yim, "Clinical usefulness of the geriatric depression scale to identify the elderly at risk of suicide," *Psychiatry Investig*, vol. 17, no. 5, pp. 481–486, May 2020, doi: 10.30773/pi.2019.0299.
- [41] Q. Chang, F. Sha, C. H. Chan, and P. S. F. Yip, "Validation of an abbreviated version of the Lubben Social Network Scale ('LSNS-6') and its associations with suicidality among older adults in China," *PLoS One*, vol. 13, no. 8, p. e0201612, Aug. 2018, doi: 10.1371/journal.pone.0201612.
- [42] Z. Ivan Santini, A. Koyanagi, S. Tyrovolas, and J. M. Haro, "The association of relationship quality and social networks with depression, anxiety, and suicidal ideation among older married adults: Findings from a cross-sectional analysis of the Irish Longitudinal Study on Ageing (TILDA)," *J Affect Disord*, vol. 179, pp. 134–141, Jul. 2015, doi: 10.1016/j.jad.2015.03.015.
- [43] L. Sweeney, C. Owens, and K. Malone, "Communication and interpretation of emotional distress within the friendships of young Irish men prior to suicide: a qualitative study," *Health Soc Care Community*, vol. 23, no. 2, pp. 150–158, Mar. 2015, doi: 10.1111/hsc.12124.
- [44] J. Cerel, J. R. Jordan, and P. R. Duberstein, "The impact of suicide on the family," *Crisis*, vol. 29, no. 1. pp. 38–44, 2008. doi: 10.1027/0227-5910.29.1.38.
- [45] V. M. B. Silenzio, P. R. Duberstein, W. Tang, N. Lu, X. Tu, and C. M. Homan, "Connecting the invisible dots: Reaching lesbian, gay, and bisexual adolescents and

- young adults at risk for suicide through online social networks,” *Soc Sci Med*, vol. 69, no. 3, pp. 469–474, Aug. 2009, doi: 10.1016/j.socscimed.2009.05.029.
- [46] N. Slesnick, J. Zhang, and L. Walsh, “Youth Experiencing Homelessness with Suicidal Ideation: Understanding Risk Associated with Peer and Family Social Networks,” *Community Mental Health Journal* 2020 57:1, vol. 57, no. 1, pp. 128–135, Apr. 2020, doi: 10.1007/S10597-020-00622-7.
- [47] R. Goldstein and D. Black, “The prediction of suicide: sensitivity, specificity, and predictive value of a multivariate model applied to suicide among 1906 patients with affective disorders,” *Archives of general*, 1991.
- [48] D. H. Hughes, “Can the Clinician Predict Suicide?,” *Psychiatric Services*, vol. 46, no. 5, pp. 449–451, May 1995, doi: 10.1176/ps.46.5.449.
- [49] M. Large and C. Ryan, “Suicide risk assessment: myth and reality,” *Int J Clin Pract*, vol. 68, no. 6, pp. 679–681, Jun. 2014, doi: 10.1111/ijcp.12378.
- [50] M. M. Large and O. Nielssen, “Suicidal Ideation and Later Suicide,” *American Journal of Psychiatry*, vol. 169, no. 6, pp. 662–662, Jun. 2012, doi: 10.1176/appi.ajp.2012.11111674.
- [51] J. Paris, “Predicting and preventing suicide: do we know enough to do either?,” *Harv Rev Psychiatry*, vol. 14, no. 5, pp. 233–240, 2006.
- [52] A. Beck, R. Beck, and M. Kovacs, “Classification of suicidal behaviors: I. Quantifying intent and medical lethality.,” *Am J Psychiatry*, vol. 132, no. 3, pp. 285–287, 1975, doi: 10.1176/appi.books.9780890425596.
- [53] A. Beck, M. Kovacs, and A. Weissman, “Assessment of suicidal intention: the Scale for Suicide Ideation.,” *Journal of consulting and*, vol. 47, no. 2, pp. 343–352, 1979, doi: <http://dx.doi.org/10.1037/0022-006X.47.2.343>.
- [54] M. E. Larsen *et al.*, “The use of technology in Suicide Prevention,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2015, pp. 7316–7319. doi: 10.1109/EMBC.2015.7320081.
- [55] J. Pestian, M. Sorter, and B. Connolly, “A machine learning approach to identifying the thought markers of suicidal subjects: a prospective multicenter trial,” *Suicide and life-*, 2017.
- [56] A. Abbasi, A. Hassan, and M. Dhar, “Benchmarking Twitter Sentiment Analysis Tools Sentiment Analysis on Social Media View project Signal Fusion and Semantic Similarity Evaluation for Social Media Based Adverse Drug Event Detection View project Benchmarking Twitter Sentiment Analysis Tools,” 2014. Accessed: Mar. 28, 2020. [Online]. Available: <https://www.researchgate.net/publication/273000042>
- [57] N. Asghar, “Automatic Extraction of Causal Relations from Natural Language Texts: A Comprehensive Survey,” May 2016, Accessed: Mar. 30, 2020. [Online]. Available: <http://arxiv.org/abs/1605.07895>
- [58] S. v. Cole, M. D. Royal, M. G. Valtorta, M. N. Huhns, and J. B. Bowles, “A lightweight tool for automatically extracting causal relationships from text,” in *Conference Proceedings - IEEE SOUTHEASTCON*, 2006, vol. 2006, pp. 125–129. doi: 10.1109/second.2006.1629336.
- [59] R. Girju and D. Moldovan, “Text Mining for Causal Relations,” 2002. Accessed: Mar. 30, 2020. [Online]. Available: www.aaai.org
- [60] H. Schwartz and L. Ungar, “Data-driven content analysis of social media: a systematic overview of automated methods,” *The ANNALS of the American*, 2015.

- [61] M. C. Monuteaux and C. Stamoulis, "Machine Learning: A Primer for Child Psychiatrists," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 55, no. 10. Elsevier Inc., pp. 835–836, Oct. 01, 2016. doi: 10.1016/j.jaac.2016.07.766.
- [62] R. Lacson and R. Khorasani, "Natural language processing: The basics (Part 1)," *Journal of the American College of Radiology*, vol. 8, no. 6, pp. 436–437, Jun. 2011, doi: 10.1016/j.jacr.2011.04.020.
- [63] M. Birjali, A. Beni-Hssane, and M. Erritali, "Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Sentiment Prediction in Social Networks," *Procedia Comput Sci*, vol. 113, pp. 65–72, 2017, doi: 10.1016/j.procs.2017.08.290.
- [64] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent Convolutional Neural Networks for Text Classification." Accessed: Apr. 01, 2020. [Online]. Available: www.aaii.org
- [65] B. Pang, L. L.-F. and T. in Information, and undefined 2008, "Opinion mining and sentiment analysis," *Foundations and Trends{texregistered} in Information Retrieval*, vol. 2, pp. 1–135, 2008, Accessed: Aug. 12, 2019. [Online]. Available: <http://www.nowpublishers.com/article/Details/INR-011>
- [66] B. Pang, L. L.-P. of the 42nd annual meeting on, and undefined 2004, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," *dl.acm.org*, Accessed: Apr. 01, 2020. [Online]. Available: <https://dl.acm.org/citation.cfm?id=1218990>
- [67] C. C. Chen and Y. De Tseng, "Quality evaluation of product reviews using an information quality framework," *Decis Support Syst*, vol. 50, no. 4, pp. 755–768, Mar. 2011, doi: 10.1016/j.dss.2010.08.023.
- [68] N. Fielding, R. Lee, and G. Blank, *The SAGE handbook of online research methods*. Sage, 2008.
- [69] M. Barnes, "User-Generated Metadata in Social Software: An Analysis of Findability in Content Tagging and Recommender Systems," 2007.
- [70] M. Lieberman, "Visualizing Big Data: Social Network Analysis," in *Digital research conference*, 2014, pp. 1–23.
- [71] Torruella Joaquim and J. Llisterri, "Diseño de corpus textuales y orales," in *Filología e informática. Nuevas tecnologías en los estudios filológicos*, 1999, pp. 45–77. Accessed: Apr. 11, 2020. [Online]. Available: http://liceu.uab.es/~joaquim/publicacions/Torruella_Llisterri_99.pdf
- [72] S. Bhattacharya *et al.*, "Developing a Multilingual Annotated Corpus of Misogyny and Aggression," Mar. 2020, Accessed: Oct. 25, 2020. [Online]. Available: <http://arxiv.org/abs/2003.07428>
- [73] L. Bentivogli and E. Pianta, "Exploiting parallel texts in the creation of multilingual semantically annotated resources: The MultiSemCor Corpus," *Nat Lang Eng*, vol. 11, no. 3, pp. 247–261, Sep. 2005, doi: 10.1017/S1351324905003839.
- [74] T. McEnery, "Corpus Linguistics," in *The Oxford Handbook of Computational Linguistics*, vol. 9780199276349, Oxford University Press, 2012. doi: 10.1093/oxfordhb/9780199276349.013.0024.
- [75] A. Ben-David, "Comparison of classification accuracy using Cohen's Weighted Kappa," *Expert Syst Appl*, vol. 34, no. 2, pp. 825–832, Feb. 2008, doi: 10.1016/j.eswa.2006.10.022.
- [76] L. Canales, C. Strapparava, ... E. B.-... C. on D., and undefined 2016, "Exploiting a bootstrapping approach for automatic annotation of emotions in texts,"

ieeexplore.ieee.org, Accessed: Apr. 02, 2020. [Online]. Available:
<https://ieeexplore.ieee.org/abstract/document/7796960/>

- [77] K. Fu, Q. Cheng, P. W. C. Wong, and P. S. F. Yip, "Responses to a Self-Presented Suicide Attempt in Social Media," *http://dx.doi.org/10.1027/0227-5910/a000221*, Jan. 2013, doi: 10.1027/0227-5910/A000221.
- [78] K. A. Hallgren, "Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial," *Tutor Quant Methods Psychol*, vol. 8, no. 1, pp. 23–34, Feb. 2012, doi: 10.20982/tqmp.08.1.p023.
- [79] F. Tapia, C. Aguinaga, and R. Luje, "Detection of Behavior Patterns through Social Networks like Twitter, using Data Mining techniques as a method to detect Cyberbullying," in *2018 7th International Conference On Software Process Improvement (CIMPS)*, Oct. 2018, pp. 111–118. doi: 10.1109/CIMPS.2018.8625625.
- [80] S. M. Vieira, U. Kaymak, and J. M. C. Sousa, "Cohen's kappa coefficient as a performance measure for feature selection," in *2010 IEEE World Congress on Computational Intelligence, WCCI 2010*, 2010. doi: 10.1109/FUZZY.2010.5584447.
- [81] M. J. Vioules, B. Moulahi, J. Aze, and S. Bringay, "Detection of suicide-related posts in Twitter data streams," *IBM J Res Dev*, vol. 62, no. 1, Jan. 2018, doi: 10.1147/JRD.2017.2768678.
- [82] X. Alameda-Pineda *et al.*, "RAVEL: An annotated corpus for training robots with audiovisual abilities," *Journal on Multimodal User Interfaces*, vol. 7, no. 1–2, pp. 79–91, Sep. 2013, doi: 10.1007/s12193-012-0111-y.
- [83] K. Bontcheva *et al.*, "GATE Teamware: A web-based, collaborative text annotation framework," *Lang Resour Eval*, vol. 47, no. 4, pp. 1007–1029, Feb. 2013, doi: 10.1007/s10579-013-9215-6.
- [84] M. Karimzadeh and A. MacEachren, "GeoAnnotator: A Collaborative Semi-Automatic Platform for Constructing Geo-Annotated Text Corpora," *ISPRS Int J Geoinf*, vol. 8, no. 4, p. 161, Mar. 2019, doi: 10.3390/ijgi8040161.
- [85] W. Ling, L. Marujo, C. Dyer, A. W. Black, and I. Trancoso, "Mining parallel corpora from Sina Weibo and Twitter," *Computational Linguistics*, vol. 42, no. 2, pp. 307–343, Jun. 2016, doi: 10.1162/COLI_a_00249.
- [86] J. Jashinsky *et al.*, "Tracking Suicide Risk Factors Through Twitter in the US," *Crisis*, vol. 35, no. 1, pp. 51–59, Jan. 2014, doi: 10.1027/0227-5910/a000234.
- [87] D. Mowery *et al.*, "Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study.," *J Med Internet Res*, vol. 19, no. 2, p. e48, Feb. 2017, doi: 10.2196/jmir.6895.
- [88] M. Purver and S. Battersby, "Experimenting with Distant Supervision for Emotion Classification," *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 482–491, 2012.
- [89] T. Wu, D. S. Weld, and J. Heer, "Local Decision Pitfalls in Interactive Machine Learning," *ACM Transactions on Computer-Human Interaction*, vol. 26, no. 4, pp. 1–27, Jun. 2019, doi: 10.1145/3319616.
- [90] X. Huang, L. Zhang, D. Chiu, T. Liu, X. Li, and T. Zhu, "Detecting Suicidal Ideation in Chinese Microblogs with Psychological Lexicons," in *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*, 2014, pp. 844–849. doi: 10.1109/UIC-ATC-ScalCom.2014.48.

- [91] A. Zhang, A. Li, and T. Zhu, "Recognizing Temporal Linguistic Expression Pattern of Individual with Suicide Risk on Social Media," *arXiv preprint arXiv:1511.09173*, 2015.
- [92] B. Desmet and V. Hoste, "Online suicide prevention through optimised text classification," *Inf Sci (N Y)*, vol. 439–440, pp. 61–78, May 2018, doi: 10.1016/J.INS.2018.02.014.
- [93] Reddit, "Deep Web," *reddit*. <https://www.reddit.com/r/deepweb/> (accessed Apr. 16, 2019).
- [94] L. Guan, B. Hao, Q. Cheng, P. Yip, and T. Zhu, "Identifying Chinese microblog users with high suicide probability using internet-based profile and linguistic features: Classification model," *JMIR Ment Health*, 2015.
- [95] S. Z. Cremades, J. M. Gomez Soriano, and B. Navarro-Colorado, "Design, compilation and annotation of a corpus for the detection of suicide messages in social networks," in *Procesamiento de Lenguaje Natural*, 2017, vol. 59, pp. 65–72.
- [96] Gómez José, "Life! Prevención del Suicidio en las Redes Sociales | Grupo de Procesamiento del Lenguaje y Sistemas de Información," 2017. <https://www.fbbva.es/equipo/analisis-sentimientos-redes-sociales-prevencion-suicidio/> (accessed Nov. 13, 2017).
- [97] A. V. Gundlapalli *et al.*, "Validating a strategy for psychosocial phenotyping using a large corpus of clinical text," *Journal of the American Medical Informatics Association*, vol. 20, no. E2, pp. e355–e364, Dec. 2013, doi: 10.1136/amiajnl-2013-001946.
- [98] S. Scheible, R. Whitt, M. Durrell, P. B.-P. of the 5th, and undefined 2011, "A gold standard corpus of Early Modern German," *dl.acm.org*.
- [99] W. P. Amorim, E. C. Tetila, H. Pistori, and J. P. Papa, "Semi-supervised learning with convolutional neural networks for UAV images automatic recognition," *Comput Electron Agric*, vol. 164, p. 104932, Sep. 2019, doi: 10.1016/j.compag.2019.104932.
- [100] J. Du, J. Xu, H. Y. Song, and C. Tao, "Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data," *BMC Med Inform Decis Mak*, vol. 17, no. 2, pp. 63–70, Jul. 2017, doi: 10.1186/s12911-017-0469-6.
- [101] A. Halike, K. Abiderexiti, and T. Yibulayin, "Semi-Automatic Corpus Expansion and Extraction of Uyghur-Named Entities and Relations Based on a Hybrid Method," *Information*, vol. 11, no. 1, p. 31, Jan. 2020, doi: 10.3390/info11010031.
- [102] K. O'Connor, A. Sarker, J. Perrone, and G. Gonzalez Hernandez, "Promoting Reproducible Research for Characterizing Nonmedical Use of Medications Through Data Annotation: Description of a Twitter Corpus and Guidelines," *J Med Internet Res*, vol. 22, no. 2, p. e15861, Feb. 2020, doi: 10.2196/15861.
- [103] C. Van Hee *et al.*, "Automatic detection of cyberbullying in social media text," *PLoS One*, vol. 13, no. 10, Oct. 2018, doi: 10.1371/journal.pone.0203794.
- [104] N. Silveira *et al.*, "A Gold Standard Dependency Corpus for English."
- [105] N. Nguyen, R. Gabud, and S. Ananiadou, "COPIOUS: A gold standard corpus of named entities towards extracting species occurrence from biodiversity literature," *Biodivers Data J*, vol. 7, no. 7, Jan. 2019, doi: 10.3897/bdj.7.e29626.
- [106] OMS, "Cada 40 segundos se suicida una persona." <https://www.who.int/es/news/item/09-09-2019-suicide-one-person-dies-every-40-seconds> (accessed Feb. 05, 2021).

- [107] J. Pestian and J. Grupp-Phelan, "A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department," *Suicide and life-*, 2016, Accessed: Aug. 27, 2017. [Online]. Available: <http://onlinelibrary.wiley.com/doi/10.1111/sltb.12180/full>
- [108] J. M. Gómez, "Language Technologies for Suicide Prevention in Social Media," *Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)*, pp. 21–29, 2014, doi: 10.3115/v1/W14-6903.
- [109] Centers for Disease Control and Prevention and others, "Youth risk behavior survey-- data summary & trends report: 2007--2017. 2018," 2018.
- [110] Datosmacro, "Suicidios 2018 | datosmacro.com," *Datosmacro*. <https://datosmacro.expansion.com/demografia/mortalidad/causas-muerte/suicidio> (accessed Oct. 28, 2020).
- [111] C. E. Osgood and E. G. Walker, "Motivation and language behavior: a content analysis of suicide notes," *J Abnorm Psychol*, vol. 59, no. American Psychological Association, pp. 58–67, Jul. 1959.
- [112] R. W. Acuña Caicedo, J. M. Gómez Soriano, and H. A. Melgar Sasieta, "Assessment of supervised classifiers for the task of detecting messages with suicidal ideation," *Heliyon*, vol. 6, no. 8, p. e04412, Aug. 2020, doi: 10.1016/j.heliyon.2020.e04412.
- [113] M. Liakata, J.-H. Kim, S. Saha, J. Hastings, and D. Rebolz-Schuhmann, "Three Hybrid Classifiers for the Detection of Emotions in Suicide Notes," *Biomed Inform Insights*, vol. 5s1, p. BII.S8967, Jan. 2012, doi: 10.4137/BII.S8967.
- [114] G. Owen, J. Belam, H. Lambert, J. Donovan, F. Rapport, and C. Owens, "Suicide communication events: Lay interpretation of the communication of suicidal ideation and intent," *Soc Sci Med*, vol. 75, no. 2, pp. 419–428, Jul. 2012, doi: 10.1016/J.SOCSCIMED.2012.02.058.
- [115] J. Pestian, H. Nasrallah, P. Matykiewicz, A. Bennett, and A. Leenaars, "Suicide Note Classification Using Natural Language Processing: A Content Analysis.," *Biomed Inform Insights*, vol. 2010, no. 3, pp. 19–28, Aug. 2010, Accessed: Jun. 16, 2018. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21643548>
- [116] J. Pestian *et al.*, "A Machine Learning Approach to Identifying the Thought Markers of Suicidal Subjects: A Prospective Multicenter Trial," *Suicide Life Threat Behav*, vol. 47, no. 1, pp. 112–121, Feb. 2017, doi: 10.1111/sltb.12312.
- [117] A. M. Schoene and N. Dethlefs, "Automatic Identification of Suicide Notes from Linguistic and Sentiment Features," in *Proceedings of the 10th SIGHUM Workshop on Language Technology for Cultural Herita*, 2016, pp. 128–133. Accessed: Jun. 13, 2018. [Online]. Available: <http://www.aclweb.org/anthology/W16-2116>
- [118] J. J. Shapero, "THE LANGUAGE OF SUICIDE NOTES," University of Birmingham, 2011. Accessed: Apr. 29, 2018. [Online]. Available: <http://etheses.bham.ac.uk/1525/1/Shapero11PhD.pdf>
- [119] M. A. Al-Garadi, K. D. Varathan, and S. D. Ravana, "Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network," *Comput Human Behav*, vol. 63, pp. 433–443, Oct. 2016, doi: 10.1016/j.chb.2016.05.051.
- [120] D. BIBER, S. CONRAD, and R. REPPEN, "Corpus-based Approaches to Issues in Applied Linguistics," *Appl Linguist*, vol. 15, no. 2, pp. 169–189, Jun. 1994, doi: 10.1093/applin/15.2.169.

- [121] J. Cohen, "A Coefficient of Agreement for Nominal Scales," *Educ Psychol Meas*, vol. 20, no. 1, pp. 37–46, Apr. 1960, doi: 10.1177/001316446002000104.
- [122] M. L. McHugh, "Interrater reliability: the kappa statistic," *Biochem Med (Zagreb)*, vol. 22, no. 3, pp. 276–282, Oct. 2012, doi: 10.11613/BM.2012.031.
- [123] A. Viera, J. G.-F. med, and undefined 2005, "Understanding interobserver agreement: the kappa statistic," 2005.
- [124] J. Carletta, "Assessing agreement on classification tasks: the kappa statistic," 1996.
- [125] Statista, "• Los idiomas más hablados en el mundo en 2020 | Statista." <https://es.statista.com/estadisticas/635631/los-idiomas-mas-hablados-en-el-mundo/> (accessed Jan. 08, 2021).
- [126] P. Koehn, "Europarl: A Parallel Corpus for Statistical Machine Translation." [Online]. Available: <http://www.europarl.eu.int/>
- [127] I. Doval, "The PaGeS Corpus, a Parallel Corpus of the Contemporary German and Spanish Language/Das PaGeS-Korpus, ein Parallelkorpus der deutschen und spanischen Gegenwartssprache/El corpus PaGeS, un corpus paralelo de textos alemanes y españoles contemporáneos," *Revista de Filología; a Alemana*, vol. 26, pp. 181–198, Jan. 2018, Accessed: Oct. 03, 2020. [Online]. Available: <https://go.gale.com/ps/i.do?p=IFME&sw=w&issn=19882823&v=2.1&it=r&id=GALE%7CA553628640&sid=googleScholar&linkaccess=fulltext>
- [128] M. R. Costa-jussà, P. L. Lin, and C. España-Bonet, "GeBioToolkit: Automatic Extraction of Gender-Balanced Multilingual Corpus of Wikipedia Biographies," Dec. 2019.
- [129] Parsifal, "Parsifal." <https://parsif.al/about/> (accessed Oct. 04, 2020).
- [130] Elsevier, "About Scopus - Abstract and citation database | Elsevier." https://www.elsevier.com/solutions/scopus?dgcid=RN_AGCM_Sourced_300005030 (accessed Apr. 13, 2021).
- [131] Uxía Gutiérrez Couto, "Guía de uso WOS 2017." Accessed: Dec. 03, 2020. [Online]. Available: <https://www.recursoscientificos.fecyt.es/>
- [132] Python Software Foundation, "About Python™ | Python.org." <https://www.python.org/about/> (accessed Dec. 03, 2020).
- [133] Google, "Googletrans: Free and Unlimited Google translate API for Python — Googletrans 3.0.0 documentation." <https://py-googletrans.readthedocs.io/en/latest/> (accessed Nov. 20, 2020).
- [134] R. Hernández-Sampieri and C. Torres, "Metodología de la investigación," 2018, Accessed: Apr. 05, 2022. [Online]. Available: <https://dspace.scz.ucb.edu.bo/dspace/bitstream/123456789/21401/1/11699.pdf>
- [135] C. A. Ramos-Galarza, "Los Alcances de una investigación," *CienciAmérica*, vol. 9, no. 3, pp. 1–6, Oct. 2020, doi: 10.33210/CA.V9I3.336.
- [136] Real Academia Española, "suicidio | Definición | Diccionario de la lengua española | RAE - ASALE," *Edición 23*, 2014. <https://dle.rae.es/suicidio?m=form> (accessed Jun. 11, 2022).
- [137] World Health Organization, "Suicide," *WHO*, 2019. <https://www.who.int/news-room/fact-sheets/detail/suicide> (accessed Apr. 10, 2019).
- [138] Real Academia Española, "prevención | Definición | Diccionario de la lengua española | RAE - ASALE." <https://dle.rae.es/prevenci%C3%B3n?m=form> (accessed Jun. 11, 2022).

- [139] Teléfono de la esperanza, "Asistencia al Suicida - Estamos para escucharte." <http://asistenciaalsuicida.com/> (accessed Dec. 22, 2020).
- [140] Samaritans, "Samaritans | Every life lost to suicide is a tragedy | Here to listen." <https://www.samaritans.org/> (accessed Nov. 26, 2022).
- [141] La-Barandilla, "La Barandilla | La voz de la Diversidad en Internet." <https://www.labarandilla.org/> (accessed Nov. 26, 2022).
- [142] Real Academia Española, "corpus | Definición | Diccionario de la lengua española | RAE - ASALE." <https://dle.rae.es/corpus?m=form> (accessed Jun. 11, 2022).
- [143] U. Bayram *et al.*, "Toward Suicidal Ideation Detection with Lexical Network Features and Machine Learning," *Northeast Journal of Complex Systems (NEJCS)*, vol. 4, no. 1, p. 2, Apr. 2022, doi: 10.22191/nejcs/vol4/iss1/2.
- [144] H. and L. M. Furstenu, "Semi-Supervised Semantic Role Labeling | Enhanced Reader," in *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, 2009, pp. 220–228. Accessed: Feb. 20, 2021. [Online]. Available: [moz-extension://7016f824-15b8-427e-ac3f-e7ede4e4b0db/enhanced-reader.html?openApp&pdf=https%3A%2F%2Fwww.aclweb.org%2Fanthology%2FE09-1026.pdf](https://www.aclweb.org/anthology/E09-1026.pdf)
- [145] L. Lowphansirikul, C. Polpanumas, A. T. Rutherford, and S. Nutanong, "A large English–Thai parallel corpus from the web and machine-generated text," *Language Resources and Evaluation 2021*, pp. 1–23, Mar. 2021, doi: 10.1007/S10579-021-09536-6.
- [146] Sierra and Gerardo, "Extracción de contextos definitorios en textos de especialidad a partir del reconocimiento de patrones lingüísticos," *Linguamática*, vol. 1, no. 2, pp. 13–37, 2009, doi: 10.21814/LM.1.2.38.
- [147] P. Resnik and N. A. Smith, "The Web as a Parallel Corpus." Accessed: Jan. 25, 2021. [Online]. Available: <http://www ldc.upenn.edu>
- [148] G. Leech, R. Garside, and M. Bryant, "CLAWS4: THE TAGGING OF THE BRITISH NATIONAL CORPUS."
- [149] M. Davies, "The 385+ million word Corpus of Contemporary American English (1990–2008+)," *International Journal of Corpus Linguistics*, vol. 14, no. 2, pp. 159–190, Jun. 2009, doi: 10.1075/ijcl.14.2.02dav.
- [150] L. Schubert and M. Tong, "Extracting and Evaluating General World Knowledge from the Brown Corpus."
- [151] B. G. Tabachnick, "Experimental Designs Using ANOVA Epidemiology and Neurobehaviour of FASD in South Africa View project CyGAMES View project," 2007. Accessed: Sep. 07, 2020. [Online]. Available: <https://www.researchgate.net/publication/259465542>
- [152] X. Zhao, "When to use Scott's π or Krippendorff's α , if ever?," in *The annual conference of Association for Education in Journalism and Mass Communication, St. Louis, USA*, 2011, pp. 4–38. Accessed: Apr. 12, 2020. [Online]. Available: http://repository.hkbu.edu.hk/cgi/viewcontent.cgi?article=1002&context=coms_conf
- [153] L. Màrquez, G. Escudero, D. Martínez, and G. Rigau, "Supervised Corpus-Based Methods for WSD," Springer, Dordrecht, 2007, pp. 167–216. doi: 10.1007/978-1-4020-4809-8_7.
- [154] A. Ben-David, "Comparison of classification accuracy using Cohen's Weighted Kappa," *Expert Syst Appl*, vol. 34, no. 2, pp. 825–832, Feb. 2008, doi: 10.1016/j.eswa.2006.10.022.

- [155] T. Koo, X. Carreras, and M. Collins, "Simple Semi-supervised Dependency Parsing."
- [156] C. Wang *et al.*, "VoxPopuli: A Large-Scale Multilingual Speech Corpus for Representation Learning, Semi-Supervised Learning and Interpretation," *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, pp. 993–1003, Jan. 2021, doi: 10.48550/arxiv.2101.00390.
- [157] B. Kitchenham *et al.*, "Systematic literature reviews in software engineering-A tertiary study," *Information and Software Technology*, vol. 52, no. 8. Elsevier B.V., pp. 792–805, Aug. 01, 2010. doi: 10.1016/j.infsof.2010.03.006.
- [158] D. and T. D. Denyer, "Producing a systematic review. - PsycNET," *Sage Publications Ltd*, 2009. <https://psycnet.apa.org/record/2010-00924-039> (accessed Jan. 12, 2021).
- [159] M. Araujo Alonso, "Las revisiones sistemáticas (I)," *Medwave*, vol. 11, no. 11, Nov. 2011, doi: 10.5867/medwave.2011.11.5220.
- [160] Parsifal Ltd., "Parsifal," www.parsifal.com, 2014. <https://parsifal.com>
- [161] J. F. Burnham, "Scopus database: A review," *Biomedical Digital Libraries*, vol. 3, no. 1. BioMed Central, p. 1, Mar. 08, 2006. doi: 10.1186/1742-5581-3-1.
- [162] Clarivate Analytics, "Trusted publisher-independent citation database - Web of Science Group." <https://clarivate.com/webofsciencelibrary/solutions/web-of-science/> (accessed Oct. 05, 2020).
- [163] Google Inc., "About Google Scholar." <https://scholar.google.com/intl/es/scholar/about.html> (accessed Oct. 05, 2020).
- [164] H. Zaugg, R. E. West, I. Tateishi, and D. L. Randall, "Mendeley: Creating communities of scholarly inquiry through research collaboration," *TechTrends*, vol. 55, no. 1, pp. 32–36, Jan. 2011, doi: 10.1007/s11528-011-0467-y.
- [165] B. Kitchenham, B. Kitchenham, and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," 2007.
- [166] B. Kitchenham, S. C.- Engineering, and undefined 2007, "Guidelines for performing systematic literature reviews in software engineering version 2.3".
- [167] C. M. D. C. Santos, C. A. D. M. Pimenta, and M. R. C. Nobre, "A estratégia PICO para a construção da pergunta de pesquisa e busca de evidências," *Revista Latino-Americana de Enfermagem*, vol. 15, no. 3. Associação Médica Brasileira, pp. 508–511, 2007. doi: 10.1590/S0104-11692007000300023.
- [168] Cristina Mamédio da Costa Santos; Cibele Andrucio de Mattos Pimental; and Moacyr Roberto Cuce Nobre, "The PICO strategy for the research question construction and evidence search," *Rev Lat Am Enfermagem*, vol. 15, no. 3, 2007.
- [169] Martínez Díaz, Juan Daniel*; Ortega Chacón, Verónica, Muñoz Ronda, and Francisco José, "El diseño de preguntas clínicas en la práctica basada en la evidencia: modelos de formulación," *Enfermería Global*, vol. 15, no. 43, Jul. 2016.
- [170] A. V. Aho and M. J. Corasick, "Efficient String Matching: An Aid to Bibliographic Search," *Commun ACM*, vol. 18, no. 6, pp. 333–340, Jun. 1975, doi: 10.1145/360825.360855.
- [171] T. Meline, "T Selecting Studies for Systematic Review: Inclusion and Exclusion Criteria," 2006.
- [172] P. C. Abrami, P. A. Cohen, and S. d'Apollonia, "Implementation Problems in Meta-Analysis," *Rev Educ Res*, vol. 58, no. 2, pp. 151–179, Jun. 1988, doi: 10.3102/00346543058002151.

- [173] J. Parraga-Alava, R. Acuna-Caicedo, J. M. Gomez, and M. Inostroza-Ponta, "An Unsupervised Learning Approach for Automatically to Categorize Potential Suicide Messages in Social Media," in *2019 38th International Conference of the Chilean Computer Science Society (SCCC)*, Jan. 2019, pp. 1–8. doi: 10.1109/sccc49216.2019.8966443.
- [174] TeleGeography, "Submarine Cable Map." <https://www.submarinecablemap.com/> (accessed Dec. 17, 2020).
- [175] SpaceX, "Starlink." <https://www.starlink.com/> (accessed Dec. 17, 2020).
- [176] HughesNet, "High-Speed Satellite Internet from HughesNet® | 844-737-2700." <https://www.hughesnet.com/> (accessed Dec. 17, 2020).
- [177] TeleGeography, "Cloud Infrastructure Map." <https://www.cloudinfrastructuremap.com/> (accessed Dec. 17, 2020).
- [178] A. Sarker and G. Gonzalez, "Portable automatic text classification for adverse drug reaction detection via multi-corpus training," *J Biomed Inform*, vol. 53, pp. 196–207, Feb. 2015, doi: 10.1016/j.jbi.2014.11.002.
- [179] J. A. T. Arandes, "El análisis de contenido como herramienta de utilidad para la realización de una investigación descriptiva. Un ejemplo de aplicación práctica utilizado para conocer las investigaciones realizadas sobre la imagen de marca de España y el efecto país de origen," *Provincia*, vol., no. 29, pp. 135–173, 2013, Accessed: Dec. 26, 2020. [Online]. Available: <http://www.redalyc.org/articulo.oa?id=55530465007>
- [180] H. D. Lasswell, "İletişim kuram ve araştırma dergisi Sayı 24 Kış-Bahar," 2007.
- [181] K. A. Neuendorf, *The Content Analysis Guidebook*. SAGE Publications, Inc, 2020. doi: 10.4135/9781071802878.
- [182] M. D. White and E. E. Marsh, "Content analysis: A flexible methodology," *Libr Trends*, vol. 55, no. 1, pp. 22–45, 2006, doi: 10.1353/lib.2006.0053.
- [183] HOLSTI and OR., "Content analysis for the social sciences and humanities," *Reading. MA : Addison-Wesley (content analysis)*, 1969, Accessed: Dec. 31, 2020. [Online]. Available: <https://ci.nii.ac.jp/naid/10007743499>
- [184] C. E. Osgood, "The representational model and relevant research materials," pp. 33–88, 1959.
- [185] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*. 1980. Accessed: Jan. 01, 2021. [Online]. Available: [https://books.google.com.ec/books?hl=es&lr=&id=nE1aDwAAQBAJ&oi=fnd&pg=PP1&dq=Krippendorff,+K.+\(1980\).+Content+analysis:+An+introduction+to+its+methodology.+Newbury+Park,+CA:+Sage.&ots=yZ7n_ohOew&sig=N7WS98k2g33rV0KRUw9cSxzJZBc#v=onepage&q&f=false](https://books.google.com.ec/books?hl=es&lr=&id=nE1aDwAAQBAJ&oi=fnd&pg=PP1&dq=Krippendorff,+K.+(1980).+Content+analysis:+An+introduction+to+its+methodology.+Newbury+Park,+CA:+Sage.&ots=yZ7n_ohOew&sig=N7WS98k2g33rV0KRUw9cSxzJZBc#v=onepage&q&f=false)
- [186] R. J. Hunt, "Percent Agreement, Pearson's Correlation, and Kappa as Measures of Inter-examiner Reliability," *J Dent Res*, vol. 65, no. 2, pp. 128–130, Feb. 1986, doi: 10.1177/00220345860650020701.
- [187] W. A. Scott, "Reliability of Content Analysis: The Case of Nominal Scale Coding on JSTOR," *Public Opin Q*, pp. 321–325, 1955, Accessed: Dec. 27, 2020. [Online]. Available: <https://www.jstor.org/stable/2746450?seq=1>
- [188] R. Benini, "Principii di Demongraphia: Manuali Barbera Di Scienze Giuridiche Sociali e Politiche (No. 29)," *Firenze, Italy: G. Barbera*, 1901.

- [189] E. Rogot and I. D. Goldberg, "A proposed index for measuring agreement in test-retest studies," *J Chronic Dis*, vol. 19, no. 9, pp. 991–1006, Sep. 1966, doi: 10.1016/0021-9681(66)90032-4.
- [190] K. L. Gwet, "Computing inter-rater reliability and its variance in the presence of high agreement," *British Journal of Mathematical and Statistical Psychology*, vol. 61, no. 1, pp. 29–48, May 2008, doi: 10.1348/000711006X126600.
- [191] W. M. Grove, N. C. Andreasen, P. McDonald-Scott, M. B. Keller, and R. W. Shapiro, "Reliability Studies of Psychiatric Diagnosis: Theory and Practice," *Arch Gen Psychiatry*, vol. 38, no. 4, pp. 408–413, Apr. 1981, doi: 10.1001/archpsyc.1981.01780290042004.
- [192] S. Janson and J. Vegelius, "On generalizations of the G index and the phi coefficient to nominal scales," *Multivariate Behav Res*, vol. 14, no. 2, pp. 255–269, Jan. 1979, doi: 10.1207/s15327906mbr1402_9.
- [193] J. Guilford, "Preparation of item scores for correlation between individuals in a Q factor analysis," in *annual convention of the Society of Multivariate Experimental Psychologists*, 1961. Accessed: Jan. 01, 2021. [Online]. Available: https://scholar.google.com/ec/scholar?hl=es&as_sdt=0%2C5&q=Guilford%2C+J.+P.+%281961%2C+November%29.+Preparation+of+item+scores+for+correlation+between+in+dividuals+in+a+Q+factor+analysis.+Paper+presented+at+the+annual+convention+of+the+Society+of+Multivariate+Experimental+Psychologists.&btnG=
- [194] J. W. Holley and J. P. Guilford, "A NOTE ON THE G INDEX OF AGREEMENT," *Educ Psychol Meas*, vol. 24, no. 4, pp. 749–753, 1964.
- [195] W. D. Perreault and L. E. Leigh, "Reliability of Nominal Data Based on Qualitative Judgments," *Journal of Marketing Research*, vol. 26, no. 2, pp. 135–148, May 1989, doi: 10.1177/002224378902600201.
- [196] R. L. Brennan and D. J. Prediger, "Coefficient kappa: Some uses, misuses, and alternatives," *Educ Psychol Meas*, vol. 41, no. 3, pp. 687–699, 1981, doi: 10.1177/001316448104100307.
- [197] T. Byrt, J. Bishop, and J. B. Carlin, "Bias, prevalence and kappa," *J Clin Epidemiol*, vol. 46, no. 5, pp. 423–429, May 1993, doi: 10.1016/0895-4356(93)90018-V.
- [198] W. J. Potter and D. Levine-Donnerstein, "Rethinking validity and reliability in content analysis," *Journal of Applied Communication Research*, vol. 27, no. 3, pp. 258–284, 1999, doi: 10.1080/00909889909365539.
- [199] L. Guttman, "The test-retest reliability of qualitative data," *Psychometrika*, vol. 11, no. 2, pp. 81–95, Jun. 1946, doi: 10.1007/BF02288925.
- [200] A. E. Maxwell, "Coefficients of agreement between observers and their interpretation," *British Journal of Psychiatry*, vol. 130, no. 1, pp. 79–83, 1977, doi: 10.1192/bjp.130.1.79.
- [201] E. M. Bennett, R. Alpert, and A. C. Goldstein, "Communications Through Limited Response Questioning," *Public Opin Q*, vol. 18, no. 3, p. 303, Autumn 1954, doi: 10.1086/266520.
- [202] K. Krippendorff, "Bivariate Agreement Coefficients for Reliability of Data," *Sociol Methodol*, vol. 2, p. 139, 1970, doi: 10.2307/270787.
- [203] L. A. Goodman and W. H. Kruskal, "Measures of Association for Cross Classifications," Springer, New York, NY, 1979, pp. 2–34. doi: 10.1007/978-1-4612-9995-0_1.
- [204] S. Siegel and N. Castellan Jr, *Nonparametric Statistics for the Behavioral Sciences*. New York, 1988.

- [205] K. L. Gwet, *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. 2014.
- [206] I. H. (Ian H.) Witten, E. Frank, M. A. (Mark A. Hall, and C. J. Pal, *Data mining : practical machine learning tools and techniques*. Morgan Kaufmann, 2010.
- [207] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, p. 10, Nov. 2009, doi: 10.1145/1656274.1656278.
- [208] R. W. Acuña Caicedo, J. M. Gómez Soriano, and H. A. Melgar Sasieta, "Bootstrapping semi-supervised annotation method for potential suicidal messages," *Internet Interventions*, vol. 28. Elsevier B.V., p. 100519, Apr. 01, 2022. doi: 10.1016/j.invent.2022.100519.
- [209] Miriadax, "Miríada X - Cursos online masivos y en abierto de forma gratuita (Massive Open Online Courses, MOOCs)," 2017. <https://miriadax.net/home?timestamp=> (accessed Sep. 30, 2017).
- [210] Edx, "edX | Cursos online gratis de Harvard, MIT y más | edX," 2021. <https://www.edx.org/es> (accessed Mar. 31, 2021).
- [211] Coursera, "Coursera | Cursos Online de Universidades Top. Unete Gratis!," 2021. <https://es.coursera.org/> (accessed Apr. 05, 2021).
- [212] P. Lison, J. Tiedemann, and M. Kouylekov, "OpenSubtitles2018: Statistical Rescoring of Sentence Alignments in Large, Noisy Parallel Corpora." Accessed: Jun. 29, 2021. [Online]. Available: <http://www.imdb.com>
- [213] D. M. Eberhard, Gary F. Simons, and Charles D. Fennig, "Ethnologue: Languages of the World," 2021. <https://www.ethnologue.com/> (accessed Mar. 20, 2021).
- [214] M. Snell-Hornby, "The Turns of Translation Studies: New paradigms or shifting viewpoints?," 2006, Accessed: Jan. 13, 2021. [Online]. Available: [https://books.google.com.ec/books?hl=es&lr=&id=I9I4t-uAJt0C&oi=fnd&pg=PP1&dq=Snell-Hornby,+M.+\(2006\)+The+Turns+of+Translation+Studies:+New+Paradigms+or+Shifting+Viewpoint%3F.+Cited+236+times.+Amsterdam:+John+Benjamins&ots=kpMOJKZpGi&sig=dzhQBxcAdF9uiQlqHL40-Rr5Dxo](https://books.google.com.ec/books?hl=es&lr=&id=I9I4t-uAJt0C&oi=fnd&pg=PP1&dq=Snell-Hornby,+M.+(2006)+The+Turns+of+Translation+Studies:+New+Paradigms+or+Shifting+Viewpoint%3F.+Cited+236+times.+Amsterdam:+John+Benjamins&ots=kpMOJKZpGi&sig=dzhQBxcAdF9uiQlqHL40-Rr5Dxo)
- [215] H. Alejandro Sánchez, "Anotación Semántica de Contenidos, descritos en Lenguaje Natural, para el Genoma Humano." https://www.researchgate.net/publication/273635557_Anotacion_Semantica_de_Contentos_descritos_en_Lenguaje_Natural_para_el_Genoma_Humano (accessed Jan. 08, 2021).
- [216] L. Abzianidze *et al.*, "The Parallel Meaning Bank: Towards a Multilingual Corpus of Translations Annotated with Compositional Meaning Representations." Accessed: Jan. 08, 2021. [Online]. Available: <http://gutenberg.org>,
- [217] S. Zhu, Y. Yang, and C. Xu, "Extracting Parallel Sentences from Nonparallel Corpora Using Parallel Hierarchical Attention Network," *Computational Intelligence and Neuroscience*, vol. 2020. Hindawi Limited, 2020. doi: 10.1155/2020/8823906.
- [218] R. Rapp, P. Zweigenbaum, and S. Sharoff, "LREC 2016 11th Workshop on Building and Using Comparable Corpora PROCEEDINGS Edited by," 2018. Accessed: Jan. 25, 2021. [Online]. Available: <https://comparable.limsi.fr/bucc2018/>
- [219] S. Jolly and R. Agrawal, "Building English–Punjabi Parallel Corpus for Machine Translation," in *Advances in Intelligent Systems and Computing*, 2021, vol. 1165, pp. 377–385. doi: 10.1007/978-981-15-5113-0_28.

- [220] L. Padró, X. Carreras, I. Chao, and M. Padró, "FreeLing: An open-source suite of language analyzers Muntsa Padro Holmes Semantic Solutions FreeLing: An Open-Source Suite of Language Analyzers," Barcelona, 2004.
- [221] G. Coppersmith, R. Leary, P. Crutchley, and A. Fine, "Natural Language Processing of Social Media as Screening for Suicide Risk," *Biomed Inform Insights*, vol. 10, p. 117822261879286, Jan. 2018, doi: 10.1177/1178222618792860.
- [222] J. Schwarzman, A. Bauman, B. Gabbe, C. Rissel, T. Shilton, and B. J. Smith, "Organizational determinants of evaluation practice in Australian prevention agencies," *Health Educ Res*, vol. 33, no. 3, pp. 243–255, Jun. 2018, doi: 10.1093/her/cyy015.
- [223] E. Gilbert, "Widespread underprovision on reddit," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 2013, pp. 803–807. doi: 10.1145/2441776.2441866.
- [224] J. Cao, T. Chen, and J. Fan, "Fast online learning algorithm for landmark recognition based on BoW framework," in *Proceedings of the 2014 9th IEEE Conference on Industrial Electronics and Applications, ICIEA 2014*, Oct. 2014, pp. 1163–1168. doi: 10.1109/ICIEA.2014.6931341.
- [225] S. Suthaharan, "Support Vector Machine," Springer, Boston, MA, 2016, pp. 207–235. doi: 10.1007/978-1-4899-7641-3_9.
- [226] G. Siglidis, C. Giatsidis, and K. Skianis, "GraKeL: A Graph Kernel Library in Python," 2020. [Online]. Available: <http://jmlr.org/papers/v21/18-370.html>.
- [227] P. Goyal, L. Behera, and T. M. McGinnity, "Application of Bayesian framework in natural language understanding," *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, vol. 25, no. 5, pp. 251–269, Sep. 2008, doi: 10.4103/0256-4602.44656.
- [228] Rami Al-Rfou, "Welcome to polyglot's documentation! — polyglot 16.07.04 documentation." <https://polyglot.readthedocs.io/en/latest/> (accessed Jun. 13, 2021).
- [229] Stack Overflow, "scikit learn - What's the difference between Sklearn F1 score 'micro' and 'weighted' for a multi class classification problem? - Data Science Stack Exchange." <https://datascience.stackexchange.com/questions/40900/whats-the-difference-between-sklearn-f1-score-micro-and-weighted-for-a-mult> (accessed May 29, 2021).
- [230] Scikit learn Developers, "Scikit learn." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html (accessed Jan. 24, 2021).