**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

EAP. DE INGENIERÍA DE SISTEMAS

**Sistema basado en el análisis de sentimientos aplicado en las redes sociales para detectar tendencias depresivas en adolescentes ¿para evitar ¿’?**

**AUTOR:**

**Joel John Regalado Morales**

1. **INTRODUCCIÓN**
   1. **Antecedentes**

La depresión es un trastorno mental común, que incluye tristeza persistente o pérdida de interés o placer acompañado de varios como sueño o apetito perturbado, sentimientos de culpa o baja autoestima, sensación de cansancio, etc. Los períodos de depresión pueden ser duraderos o pueden ir y venir durante diferentes períodos de tiempo. (WHO 2016)

Anteriormente, la teoría psicoanalítica y psicológica tradicional propuso que todos los adolescentes tenían períodos de depresión porque experimentan conflictos dentro de sí mismos comúnmente denominados "trastornos de la adolescencia". Se creía que esto era una etapa normal de desarrollo. G. Stanley Hall (1904) Creía que los adolescentes normales experimentaban oscilaciones de ánimo anchas y un funcionamiento variable, pero que esto no significaba psicopatología. Más recientemente, estudios han refutado la opinión de que la agitación de los adolescentes es normativa. Los adolescentes tienen periodos de sentirse solos, aislados de sus compañeros, conflictos de experiencia con la familia y los profesores, pero para la mayoría de los adolescentes, estos períodos no persisten durante un período significativo de tiempo o deterioran gravemente su desenvolvimiento. (Mulfon Laura 2004)

La depresión en adolescentes ha adquirido gran notoriedad en los últimos años, llegando a ser considerada como un problema de salud pública por su frecuencia creciente y por las consecuencias que derivan de esta. La comunidad científica reconoció oficialmente su existencia recién en 1970, durante el 4º Congreso de la Unión Europea de Paidopsiquiatras realizada en la ciudad de Estocolmo, abriéndose nuevos retos que fueron y deben ser investigados para su mejor compresión y adecuado tratamiento (INSM HD – HN, 2003)

Actualmente, se cuenta con pocos instrumentos clínicos válidos y fiables para evaluar el estado depresivo, uno de ellos, es la “Escala Tetra-dimensional para el seguimiento de la depresión (EDT – Alonso Fernandez)” que constituye una prueba para efectuar el screening (estrategia aplicada sobre una población para detectar una enfermedad en individuos sin signos o síntomas de esa enfermedad) de la depresión. Su aplicación permite distinguir inmediatamente dos sectores: el de los excluidos de la depresión y el de los posibles depresivos. (CET-ED 2011)

Thomas Insel, gurú en la salud mental, manifestó su incomodidad debido a que, a diferencia del SIDA y los linfomas, las enfermedades mentales todavía se diagnosticaban de manera subjetiva, de acuerdo a los conjuntos de síntomas recolectados en el DSM, en lugar de una medición objetiva de laboratorio como un biomarcador.

En 2012, GE Healthcare realizó una encuesta online para determinar si las web sites pueden ayudar a fomentar un comportamiento saludable, donde se obtuvo los siguientes resultados.

|  |  |
| --- | --- |
| **Opiniones de los consumidores** | **Porcentaje** |
| Uso de las redes sociales en los últimos 12 meses | 80% |
| Discusión sobre información de salud en línea | 26% |
| Cambiaron su comportamiento de salud como resultado | 30% |
| Usó para obtener o dar apoyo de a otras personas  con fines de salud o de acondicionamiento físico. | 35% |
| Consideran que los medios de comunicación social es una manera rápida y fácil de obtener información sobre salud o recomendaciones | 49% |
| Consideran una buena manera de obtener diferentes opiniones de una amplia gama de personas. | 47% |
| Confía en la información obtenida en internet | 45% |
| Considera exacta la información obtenida en internet | 44% |

Esta encuesta realizada a 2100 ciudadanos, demuestra que las redes sociales son una importante fuente de información que pueden ayudar a mejorar la salud de las personas; además el Internet se ha convertido en una plataforma para expresar estados de ánimo, sentimientos individuales de la vida cotidiana, donde los autores comparten sus pensamientos en blogs web, micro-blogs, foros, anuncios u otros medios.

[Tung C, 2015] Propone una técnica de minería de textos [E3] y un modelo de advertencia de tendencia a la depresión impulsado por eventos para predecir la tendencia de la depresión de los bloggers o post analizando sus artículos publicados, en la comparación con DMS, esta última, obtuvo una mejor tasa de precisión de 66.6% frente a un 62.4 % del modelo planteado, sin embargo, añadiendo más términos relacionados a la depresión, se puede cambiar la diferencia.

Existen diversas investigaciones que proponen el diseño, desarrollo y aplicación de sistemas para la detección y evaluación de la depresión en adolescentes, las cuales se detallara en el capítulo 2.

* 1. **Problemas**

El tener un episodio depresivo durante una etapa donde se define la formación personal, tiene consecuencias a largo plazo en muchos ámbitos de la vida. Las personas que padecen depresión en la adolescencia, generalmente implica tener una vida adulta con menos amistades, menos redes de apoyo, mayor estrés y menor alcance educaciones, ocupacional y económico.

Según la Organización Mundial de Salud, la depresión es un trastorno mental frecuente. Se calcula que a más de 300 millones de personas en el mundo la padecen, aunque existen tratamientos para tratar la depresión, más del 50% no recibe estos tratamientos siendo el principal obstáculo la falta de recursos y de personal sanitario capacitados, además de la estigmatización de los trastornos mentales y la evaluación clínica inexacta.

Ello es constatado en una investigación realizada en los centros de atención primaria realizada en Brasil a jóvenes entre 13 y 19 años, teniendo como resultados un 99,2 % de casos de depresión no identificados.

* 1. **Objetivos**
     1. **Objetivo General**

Implementación de un sistema software aplicando análisis de sentimientos a los datos que se encuentran en las redes sociales para detectar depresión en adolescentes.

* + 1. **Objetivo Específico**
* Revisar los métodos y técnicas empleados para el diagnóstico de la depresión.
* Elaboración de una base de reglas que permita entender las diferentes maneras en que el contexto de una publicación en Redes Sociales puede afectar el sentimiento.
* Validar y determinar la eficiencia del sistema desarrollado en comparación con el Cuestionario estructural tetra-dimensional para la depresión.
  1. **Justificación**

Más de 350 millones de personas en el mundo sufren o han sufrido depresión, y en los casos más graves las personas afectadas pueden llegar al suicidio, por ello, es un trastorno que es importante tener en cuenta, sobre todo porque existen terapias eficaces para dejar atrás esos episodios de gran perturbación.

Walter Castillo Martell, director del Instituto Nacional de Salud Mental Honorio Delgado-Hideyo Noguchi, explicó que las conductas suicidas están asociadas a la depresión y se forman desde la infancia. “*Si uno deja que la depresión avance, en muchos casos va a llegar al suicidio, aunque no en todos”*, sostuvo.

Según las estadísticas de la institución, el 80% de los 334 casos de suicidios presentados en el país el año 2015 tuvo su origen en problemas de ánimo y depresión. “De cada 20 personas mayores de 15 años, una se deprime al año; de cada 20 personas con episodio de depresión, una intenta suicidarse; y de cada 20 intentos de suicidio, uno llega a consumarse”, detalló Castillo.

Castillo menciona que el suicidio no es, como siempre se ha pensado, una reacción ante una circunstancia adversa específica que vive una persona, sino el desenlace fatal de un proceso que en la práctica es también una enfermedad. **Algunas personas nacen con la disposición genética para ser depresivos** y otros nacen más resistentes, “pero si en la infancia estas personas vulnerables sufren de abandono, descuido o poca atención, tendrán mayor predisposición a desarrollar una depresión mucho más severa”

Un 20% se da cuenta de que necesita atención, pero no hace nada al respecto. En tanto, el resto (55%) no advierte la enfermedad y no le presta atención.



* 1. **Alcances y Limitaciones**
     1. **Alcances**

El sistema propuesto ayudara en el proceso de detección de la depresión determinando la probabilidad que la persona padezca o no depresión.

El sistema podrá ser utilizado por personal con/sin experiencia en el tema.

* + 1. **Limitaciones**

El sistema solo

El sistema solo da como resultado el diagnóstico, más no el tratamiento o recomendación.

Solo se podrá tener acceso a la información de perfiles públicos y/o a perfiles con perfil privado previa autorización.

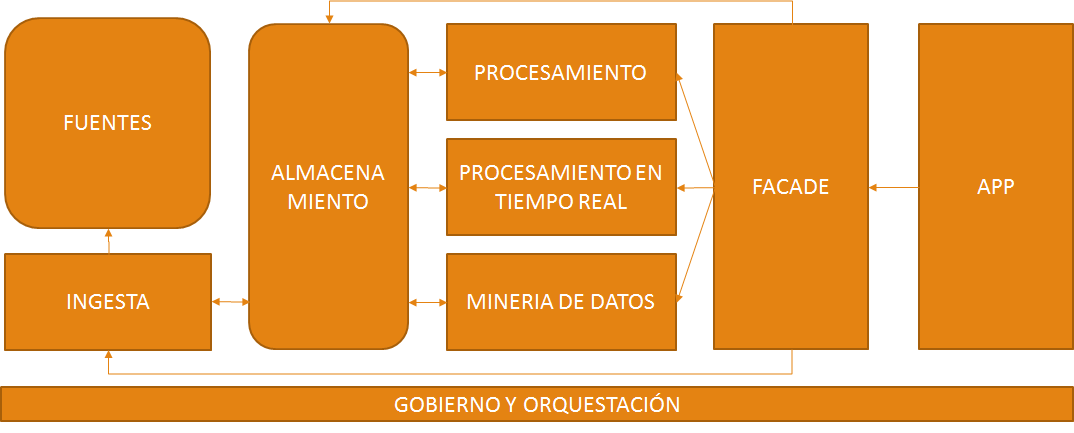
1. **ESTADO DEL ARTE**
   1. **Marco Teórico**
      1. **Big Data**

Es un marco de trabajo para el procesamiento de grandes cantidades de datos, con estructuras diferentes a grandes velocidades, ello, surge en respuesta a la creación de más y más datos en formatos de video e imagen no estructurados, que las tecnologías tradicionales de gestión de datos son inadecuadas para procesar.

Big Data tiene tres columnas sobre las cuales se sostienen todos sus otros conceptos y las implementaciones de sus herramientas:

* El volumen: Vivimos en un mundo hiper-informado, en donde la sobre-saturación de datos es parte del día a día.
* La variedad: Los datos se producen en distintos formatos (texto, audio, imágen, video, binario, raw, etc), y poseen distinta naturaleza (correo, documentos, mensajería, grabaciones, streaming, etc)
* La velocidad: La sociedad actual es impaciente, queremos respuestas lo más pronto posible

Arquitectura general de sistemas basados en big data



Donde:

* Fuentes: Repositorios que conforman la fuentes de datos del sistema
* Almacenamiento: Repositorio que almacena e integra los datos que son extraídos de las fuentes de datos.
* Ingesta: Componente que extrae los datos desde las fuentes de datos y los coloca en el repositorio final
* Procesamiento: Componente que implementa los procesos de reglas de negocio
* Procesamiento en tiempo real: Componente que implementa los procesos de reglas de negocio que deben tener un tiempo de respuesta cercano al tiempo real
* Minería de datos: Componente que implementa los modelos predictivos/descriptivos que permiten la generación de conocimiento al negocio
* Facade: Componente que implementa la interfaz de comunicación de los procesos de negocio hacia el exterior.
* App: Son los componentes que conforman el ecosistema de aplicaciones empresariales y acceden a los procesos de big data por medio del facade.
  + 1. **Análisis de Sentimientos o Minería de Opinión**

Análisis de sentimiento (también conocido como minería de opinión) se refiere al uso de procesamiento de lenguaje natural, análisis de texto y lingüística computacional para identificar y extraer información subjetiva de unos recursos. Desde el punto de vista de la minería de textos, el análisis de sentimientos es una tarea de clasificación masiva de documentos de manera automática, en función de la connotación positiva o negativa del lenguaje ocupado en el documento. Es importante mencionar que estos tratamientos generalmente "se basan en relaciones estadísticas y de asociación, no en análisis lingüístico"

El procesamiento de lenguajes naturales —abreviado PLN, o NLP del idioma inglés Natural Language Processing— es un campo de las ciencias de la computación, inteligencia artificial y lingüística que estudia las interacciones entre las computadoras y el lenguaje humano.

En términos generales, el análisis de sentimiento intenta determinar la actitud de un interlocutor o un escritor con respecto a algún tema o la polaridad contextual general de un documento. La actitud puede ser su juicio o evaluación, estado afectivo, o la intención comunicativa emocional.

En el diseño de los sistemas se combinan distintas áreas como son las técnicas de minería de datos, la selección de atributos y los algoritmos de árboles de decisión.

* El procesamiento del lenguaje natural (PLN) transforma el texto en un lenguaje que la máquina pueda entender, el Big Data recaba gran cantidad de datos para obtener un análisis más preciso al mejorar el desempeño de los algoritmos.
* Inteligencia Artificial (IA) utiliza la información que le proporciona el PLN para determinar las categorías de sentimientos y sus correspondientes polaridades: satisfacción-insatisfacción, confianza-temor, amor-odio, felicidad-tristeza.
  + 1. **Depresión**

La tristeza y la melancolía son dos sentimientos presentes en algún momento de la vida de todas las personas, al igual que la alegría y el placer. Los dos primeros no son en sí patológicos, pero en algunas ocasiones pueden llegar a serlo para ciertas personas. Cuando el estado de ánimo de un individuo en un determinado momento de su vida sufre sentimientos severos y prolongados de tristeza o síntomas relacionados que afectan a su capacidad para relacionarse con otros, trabajar o afrontar el día, la tristeza se convierte en una enfermedad, que se conoce como depresión.

Existen varios tipos de trastornos depresivos.

* El trastorno depresivo grave, también llamado depresión grave, se caracteriza por una combinación de síntomas que interfieren con la capacidad para trabajar, dormir, estudiar, comer, y disfrutar de las actividades que antes resultaban placenteras. La depresión grave incapacita a la persona y le impide desenvolverse con normalidad. Un episodio de depresión grave puede ocurrir solo una vez en el curso de vida de una persona, pero mayormente, recurre durante toda su vida.
* El trastorno bipolar, también llamado enfermedad maniaco depresiva, no es tan común como la depresión grave o la distimia. El trastorno bipolar se caracteriza por cambios ciclicos en el estado de ánimo que van desde estados de ánimo muy elevado (por ejemplo, manía) a estados de ánimo muy bajo (por ejemplo, depresión). Más información sobre el trastorno bipolar El trastorno distímico, también llamado distimia, se caracteriza por sus síntomas de larga duración (dos años o más), aunque menos graves, pueden no incapacitar a una persona pero sí impedirle desarollar una vida normal o sentirse bien. Las personas con distimia también pueden padecer uno o más episodios de depresión grave a lo largo de sus vidas. Algunas formas de trastorno depresivo muestran características levemente diferentes a las descritas anteriormente o pueden desarrollarse bajo circunstancias únicas. Sin embargo, no todos los científicos están de acuerdo en cómo caracterizar y definir estas formas de depresión. Estas incluyen: Depresión psicótica, que ocurre cuando una enfermedad depresiva grave está acompañada por alguna forma de psicosis, tal como ruptura con la realidad, alucinaciones, y delirios.
* Depresión posparto, la cual se diagnostica si una mujer que ha dado a luz recientemente sufre un episodio de depresión grave dentro del primer mes después del parto. Se calcula que del 10 al 15 por ciento de las mujeres padecen depresión posparto luego de dar a luz.1
* El trastorno afectivo estacional, se caracteriza por la aparición de una enfermedad depresiva durante los meses del invierno, cuando disminuyen las horas de luz solar. La depresión generalmente desaparece durante la primavera y el verano. El trastorno afectivo estacional puede tratarse eficazmente con terapia de luz, pero aproximadamente el 50% de las personas con trastorno afectivo estacional no responden solamente a la terapia de luz. Los medicamentos antidepresivos y la psicoterapia pueden reducir los síntomas del trastorno afectivo estacional, ya sea de forma independiente o combinados con la terapia de luz.
  + 1. **Redes Sociales**

Red de las relaciones existentes entre las personas implicadas en una sociedad,

En la actualidad, el análisis de las redes sociales se ha convertido en un método de estudio de las ciencias como la antropología o la sociología. Internet y las nuevas tecnologías favorecen el desarrollo y ampliación de las redes sociales.

La teoría de los seis grados de separación afirma que cada individuo del planeta está conectado con el resto. Esta relación se basa en una cadena de conocidos que no supera las 6 personas. Esta hipótesis ha intentado ser demostrada desde su origen a principios del siglo XX.

Las redes sociales en internet son aplicaciones web que favorecen el contacto entre individuos. Estas personas pueden conocerse previamente o hacerlo a través de la red. Contactar a través de la red puede llevar a un conocimiento directo o, incluso, la formación de nuevas parejas.

Las redes sociales en internet se basan en los vínculos que hay entre sus usuarios. Existen varios tipos de redes sociales:

* Redes sociales genéricas. Son las más numerosas y conocidas. Las más extendidas en España son Facebook, Instagram, Google+ y Twitter.
* Redes sociales profesionales. Sus miembros están relacionados laboralmente. Pueden servir para conectar compañeros o para la búsqueda de trabajo. Las más conocidas son LinkedIn, Xing y Viadeo.
* Redes sociales verticales o temáticas. Están basadas en un tema concreto. Pueden relacionar personas con el mismo hobbie, la misma actividad o el mismo rol. Las más famosas son Flickr, Pinterest y YouTube.

**REFERENCIAS BIBLIOGRAFICAS**

* [1] Prevalencia y Factores asociados con el episodio depresivo en adolescentes de la población urbana de tres ciudades en la sierra peruana 2003, INSM HD – HN 2013, Lima-Perú
* [2] World Health Organization and Columbia University. Group Interpersonal Therapy (IPT) for Depression (WHO generic field-trial version 1.0). Geneva, WHO, 2016.
* [3] Mufson Laura, Pollack, Moreau, Myrna, INTERPERSONAL PSYCHOTHERAPY FOR DEPRESSED ADOLESCENTS, The Guilfor Press 2004
* [4] APA, Manual Diagnostico y Estadístico de los Trastorno Mentales 5ta edición, Editorial Médica Panamericana, (2014), Madrid-España.
* [5] Dmytro Karamshuk, Frances Shaw: Bridging big data and qualitative methods in the social sciences: A case study of Twitter responses to high profile deaths by suicide, Only Social Networks and Media, 2017.
* [6] MIT Technology Review: La tecnología podría diagnosticar enfermedades mentales de forma objetiva Fecha de visita: (28/05/2017), https://www.technologyreview.es/s/5198/la-tecnologia-podria-diagnosticar-enfermedades-mentales-de-forma-objetiva
* [7] <http://www.insm.gob.pe/investigacion/archivos/SSMP/2012-SSMP/#/6> visitada: 22/05/2017
* [8] Franciso Alonso Fernandez: Cuestionario Estructural Tetradimensional para la Depresión, TEA Publicaciones de Psicología Aplicada 2011.
* [9] Tung C, Lu W. Analyzing depression tendency of web posts using an event-driven depression tendency warning model. Artif Intell Med (2015), <http://dx.doi.org/10.1016/j.artmed.2015.10.003>
* Organización Mundial de Salud : <http://www.who.int/mediacentre/factsheets/fs369/es/> visitada 07/06/2017
* Liu, Bing. 2012. *Sentiment analysis and opinion mining*. Synthesis Lectures on Human Language Technologies, 5(1):1-167.