

Identificando el riesgo de suicidio a través de una red neuronal capsular usando datos de Facebook

Juan D. Angarita - Sharol G. Torres
Universidad Popular del Cesar

Una de las principales causas de muerte que ha tenido un aumento dramático en los últimos años en Colombia es el suicidio. Esta situación tiene un componente inherente que lo hace difícil de prevenir: es muy complejo identificar a las personas en riesgo. En el presente estudio mostraremos cómo (usando un enfoque tecnológico) es posible identificar con un porcentaje de acierto elevado a las personas en riesgo. Utilizando la información pública de Facebook de personas que se han suicidado, hemos entrenado una red neuronal capsular para que pueda llevar a cabo dicha labor. Elementos tales como: cambios en la preferencia de colores, horas de publicación recurrente y contenido textual de sus perfiles, han podido estar relacionados con la problemática.

Keywords: suicidio, facebook, redes neuronales capsulares, análisis de sentimientos

Introducción

A día de hoy, identificar de forma masiva los factores, señales y patrones que dan origen al pensamiento suicida o la consumación del mismo sigue siendo una tarea que parece superar todos nuestros esfuerzos. Evidentemente, la conducta suicida representa un desafío complejo para las ciencias humanas, en el que las causas que desencadenan la crisis se interrelacionan de tal forma, que detectarlas puede ser un proceso muy largo. Como flagelo social, este ha sido foco investigativo en diversas disciplinas, cada una encarando el problema de una forma distinta y, sin embargo, los resultados siguen siendo poco aplicables.

Entre tanto, según estadísticas del INMLCF (2016), la autoeliminación ha venido escalando entre las principales causas de muerte por factores externos en Colombia, siendo la población más afectada sujetos entre 20 y 34 años.

Considerando que la población más afectada por esta problemática es también la de mayor acceso a redes sociales (IWS, 2017), y dado el aumento de

la tendencia a publicar contenido sobre su estado de ánimo en perfiles públicos de Internet, podemos anticipar que explorar estos espacios en busca de elementos comunes entre ellos, puede conducir a la identificación de patrones ignorados hasta ahora, lo que a su vez aumentaría la detección temprana de riesgo potencial e ideación e intento suicida.

Elementos tales como: cambios en la preferencia de colores (Carruthers, Morris, Tarrier & Whorwell, 2010), aumento de actividad nocturna (Tubbs y col., 2018) y un discurso marcado por algunas palabras clave (Sally, Margaret & Holly, 2016), han probado estar relacionados con la depresión y aparición de ideas suicidas y las redes sociales son sitios propicios para rastrear dichos elementos.

Uno de los elementos limitantes en la lucha contra el suicidio es la falta de detección temprana, es por ello que este estudio se basará en el perfilamiento de las señales y todas las variables que inciden en la consumación del suicidio y sus manifestaciones previas.

Se sostiene que la funcionalidad de las alertas

primarias debe tomarse como señales que conducirán a la necesaria detección temprana, lo que implica necesariamente el reconocimiento del suicidio como proceso en el que existen diversos factores intervinientes, que finalmente llevan al sujeto a la decisión de consumir el acto. Por lo que se considera fundamental tener conocimiento sobre el proceso y los eventos que ocurren anteriormente a que la persona llegue a esta decisión fatal. Dichos eventos son la ideación suicida y el intento de suicidio (Sánchez-Sosa, Villarreal-González, Musitu & Martínez Ferrer, 2010).

Lamentablemente, la detección temprana sigue siendo una deuda pendiente de la investigación con la sociedad. Sin embargo, tomando como base el más reciente avance de la inteligencia artificial y usando como catalizador un algoritmo de redes neuronales capsulares, analizaremos la producción textual, visual, los patrones de actividad, las preferencias en los cambios de colores y otras variables recolectadas en perfiles públicos de suicidas en facebook; con el objetivo de poder entrenar eficientemente un modelo que nos permita usando la misma información proveniente de perfiles comunes, identificar aquellas personas que están dando señales de estar ante un riesgo inminente.

Confiamos en que de la integración de algoritmos de inteligencia artificial y conocimientos psicológicos, vendrá la próxima revolución de nuestro campo. Existen numerosas variables cuantitativas que podrían darnos muestras claras de patrones que a día de hoy siguen velados para la investigación tradicional, puesto que, nuestra capacidad de detección de patrones y procesamiento numérico palidece ante la sofisticación actual de las máquinas.

Paradigma de redes neuronales actual

En los últimos años, las investigaciones en el campo de inteligencia artificial (en lo que se refiere al aprendizaje profundo) se centraron en cómo aumentar la profundidad de una red, agregando más capas para lograr un mayor grado de abstracción (comenzando desde las primeras capas con-

volucionales capaces de extraer formas pequeñas, ángulos o intensidad de color, vamos paso a paso combinando características simples en diseños cada vez más complejos)(Sabour, Frosst & Hinton, 2017).

Para hacer esto, tenemos que mantener bajo control el número de parámetros (y los tiempos computacionales) usamos un operador, común en todas las redes profundas, osea un pooling que nos permite que reduzca el número de parámetros reduciendo progresivamente el componente espacial seleccionando los valores más altos pero perdiendo la información espacial relacionada con las características extraídas(Abadi y col., 2016).

Redes neuronales capsulares

Por lo tanto, la investigación sobre la arquitectura de las redes neuronales se centra progresivamente en el objetivo de aprender a generalizar mejor en lugar de proporcionar datos cada vez más procesados (Zeiler & Fergus, 2013). Se realizó un primer paso con las redes neuronales capsulares, donde nuestro objetivo es obtener una Equivariancia (invarianza de las rototraslaciones) que reemplaza al operador de pooling con una nueva estructura denominada: Dynamic Routing.

Una cápsula es un grupo de neuronas. La actividad de un vector de una cápsula representa los parámetros de instanciación cuando un objeto (o parte de él) viene detectado por la red. La longitud de este vector representa la probabilidad de existencia de la clase dada, mientras que la orientación del vector codifica información espacial (por ejemplo, rotaciones y traducciones) dentro de una matriz de posa.

Arquitectura redes neuronales capsulares

La principal diferencia en las redes capsulares, por lo tanto, consiste en una arquitectura no profunda.

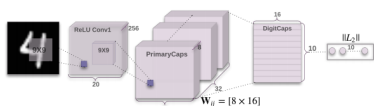


Figura 1. Esquema básico red neuronal capsular

La primera capa convolucional (256 kernels, 9x9, stride igual a uno, activación ReLu) En la primera capa convolucional vamos a convertir la intensidad de los píxeles en características de bajo nivel, no nos importa mantener la posición espacial relativa; de hecho, queremos utilizar todas esas propiedades importantes para compartir parámetros útiles para reducir los tiempos computacionales.

PrimaryCaps consiste en dos capas ocultas: La primera capa corresponde al proceso de rendering el cual lo hemos descrito antes. La segunda capa, por el contrario, es convolucional (32 cápsulas 8 dimensionales) donde cada cápsula representa una característica extraída, por lo tanto, a través de una convolución de 9x9 kernels y un ritmo igual a dos. El output de PrimaryCaps consiste por lo tanto, en cápsulas de 32x6x6 8 dimensiones.

La última capa por reproducir en forma vectorial la FC-Layer, es decir, vamos a reducir el número de neuronas (en este caso de cápsulas) para obtener una para cada clase del objetivo (Jaderberg, Simonyan, Zisserman & Kavukcuoglu, 2015).

- Reemplazar las neuronas con las cápsulas.
- Reemplazar el max-pooling con dynamic routing.
- Arquitectura no profunda.

Métodos

Participantes

Para hacer la identificación de los patrones relacionados con la conducta suicida, utilizamos 20 perfiles de Facebook con contenidos públicos de personas que cometieron, intentaron cometer o han estado en situación de riesgo frente al suicidio. El promedio de edad de los sujetos referenciados en

los perfiles es de 21 años ($\sigma = 4,5$), 11 hombres y 9 mujeres.

Materiales

Para la recolección masiva de la información se utilizó un “scraper”, tomando como base el framework Selenium, la información fue almacenada en bases de datos usando el motor Postgresql y los algoritmos de inteligencia artificial se apoyaron en los servicios cognitivos de Microsoft, Amazon web services y Tensorflow. Todos los algoritmos mencionados anteriormente fueron programados en el lenguaje Python.

Procedimiento

La recolección masiva de la información se hizo de manera autónoma y recursiva por el “scraper” que fue programado para escanear todos los perfiles y almacenar en la base de datos la información general, fotos subidas, publicaciones compartidas, estados escritos, reacciones, lista de amigos e intereses.

Posteriormente, con toda la información recolectada y almacenada en la base de datos se procedió a utilizar una red neuronal capsular para examinar todas las publicaciones e imágenes compartidas con el propósito de identificar los elementos en cada imagen, su esquema de colores y la presencia de elementos textuales.

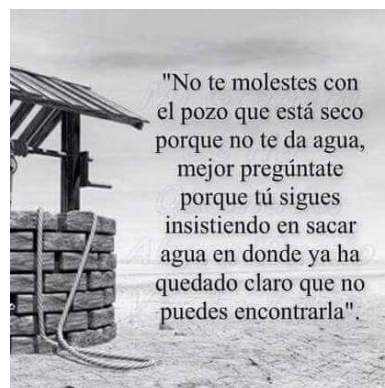


Figura 2. Una de las imágenes compartidas

La figura 1 muestra una de las imágenes compartidas en el perfil de uno de los suicidas. Para

dicha imagen el algoritmo extrajo el texto para su posterior análisis y generó la siguiente tabla de colores:

- Color 1: Plateado(196, 196, 196). Píxeles: 153398. Porcentaje: 86.75 %
- Color 2: Gris oscuro(111, 110, 116). Píxeles: 16589. Porcentaje: 9.38 %
- Color 3: Negro(9, 9, 11). Píxeles: 6833. Porcentaje: 3.86 %

Cuando todas las imágenes se procesaron, se prosiguió a tomar el texto escrito en cada imagen, publicación y elemento compartido para evaluarlo a través de un analizador de sentimientos. El enfoque utilizado fue dividir los textos en frases y examinar una por una en los servicios cognitivos de Microsoft. Una vez obtenido el puntaje de cada frase (entre 0 y 1), se examinó el texto completo y se compararon los resultados para evitar que la presencia de algunas frases positivas alteraran el puntaje de una publicación mayoritariamente negativa.

Resultados

De entre los perfiles que fueron sometidos al análisis, se encontraron que 5 de ellos no contenían la cantidad mínima de publicaciones en el último mes antes de la fecha de suicidio, por lo tanto, fueron descartados para evitar que dicha limitación afectara los resultados finales. Esta es una dificultad que responde a dos causas principales: la primera de ellas es la política de privacidad de datos de Facebook, que será discutida más adelante y la segunda es la tendencia que tienen algunas personas a alejarse del ojo público semanas antes de cometer el acto (Ursano y col., 2015).

Por otro lado, al hacer el análisis estadístico de los valores arrojados por los perfiles restantes, la inteligencia artificial nos permitió observar tres patrones importantes que en el pasado han probado estar relacionados con la ideación suicida y el acto

en sí mismo. En primer lugar, la figura 2 nos muestra los porcentajes de representatividad que tuvieron los colores en las imágenes compartidas por los sujetos en el último mes anterior a su muerte.

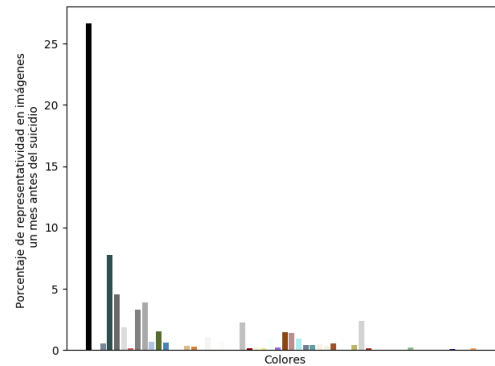


Figura 3. Imagen acumulativa de los porcentajes de representatividad de los perfiles acumulados

La figura nos muestra que existe una preferencia marcada por los sujetos en riesgo de suicidio hacia cierta gama de colores. Los colores más sobresalientes son el negro, las diferentes tonalidades de grises y algunos otros colores como el verde, rojo y café en sus tonos más oscuros. Estos datos coinciden con el estudio llevado a cabo por Carruthers y col. (2010), cuyos datos demostraban que las personas en estado de depresión o con problemas de ansiedad tienen una tendencia a preferir los colores en la escala de grises y los tonos oscuros de entre los 38 colores estudiados.

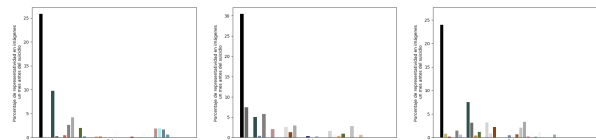


Figura 4. Pre- Figura 5. Pre- Figura 6. Preferencia de co- ferencia de co- ferencia de colores - Sujeto 1 lores - Sujeto 4 lores - Sujeto 5

Las figuras 3, 4 y 5 nos muestran el caso puntual de tres de los sujetos con mayor número de publicación de imágenes y cuya tendencia es muy notable. Cabe aclarar, para cerrar este patrón, que el porcentaje desbordado del color negro puede estar relacionado con el hecho de que muchas de

las imágenes que contenían textos, estaban escritas con fuentes de dicho color.

El segundo patrón observado, relaciona el análisis sentimental hecho al contenido textual de las publicaciones, el texto encontrado en las imágenes y el texto compartido proveniente de publicaciones de terceros que los sujetos publicaban en sus perfiles (Braithwaite, Giraud-Carrier, West, Barnes & Hanson, 2016).

En este caso en particular pudimos observar que es posible que existan dos patrones relacionados al tono de los textos y que eso implicaría que habría que entrenar las herramientas para poder generar alertas adecuadas para las dos situaciones (Sally y col., 2016). Uno de ellos hace referencia a aquellos sujetos cuya disminución en el puntaje se presenta de forma gradual, mientras que parecen existir algunos sujetos cuyo deterioro es muy rápido.

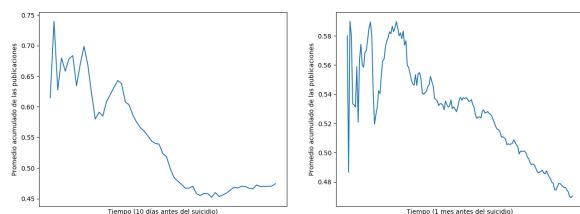


Figura 7. Análisis textual - Sujeto 7

Figura 8. Análisis textual - Sujeto 13

Como podemos observar en las figuras 6 y 7, la escala de tiempo del sujeto 7 es de tan sólo diez días, mientras que el sujeto 13 tiene una escala de un mes. Esto se debe a que la inteligencia artificial hacía el corte de datos cuando encontraba que el sujeto regresaba a los valores promedio. La figura 8, por otro lado, muestra los dos sujetos en una línea de tiempo de 50 días con el objetivo de que se aprecie con mayor claridad la aparición y diferenciación de los dos patrones mencionados anteriormente.

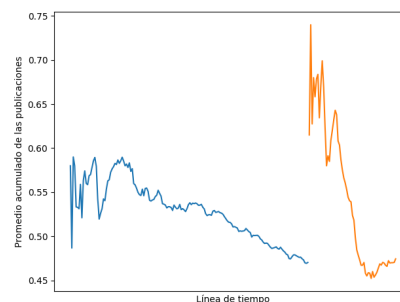


Figura 9. Análisis textual - Sujetos 7 y 13

Un posible patrón interesante que nos gustaría medir en el futuro y para el que necesitaríamos un número mayor de datos es el que relaciona los dos patrones anteriores con el estado de ánimo general de los sujetos. A priori, parece observarse que aquellas personas cuyos estados de ánimo están por debajo del promedio en condiciones normales, suelen tener la disminución gradual, mientras que aquellas personas que mantienen un estado de ánimo en la media o superior parecen ser más susceptibles a las disminuciones repentinas.

El último patrón está relacionado con la aparición o aumento de la actividad nocturna por parte de los sujetos antes del suicidio. Puntualmente, hemos medido la actividad realizada entre las 12 de la noche y las 6 de la mañana. Esta franja ha sido identificada por múltiples estudios como la más común para la ejecución del suicidio y la que mayor aumento en los niveles de actividad presenta en los días anteriores al acto. (Tubbs y col., 2018) (Littlewood, Gooding, Kyle, Pratt & Peters, 2016) (Koyawala, Stevens, McBee-Strayer, Cannon & Bridge, 2015) (Perlis y col., 2016).

En promedio, la actividad aumentó entre un 40 y un 60 % ($\sigma = 12,5\%$) durante los días previos a la consumación del suicidio.

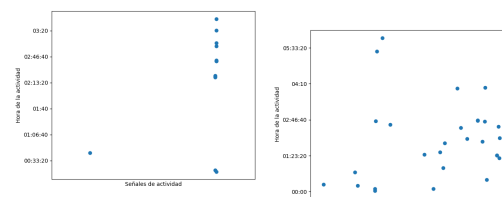


Figura 10. Act. nocturna - Sujeto 14

Figura 11. Act. nocturna - Sujeto 9

La figura 9 nos muestra la aparición repentina de actividad nocturna en el sujeto 14, cuya actividad nocturna previa era prácticamente nula. Personas con un patrón de actividad nocturna alto (figura 10), también vieron incrementado el nivel de dicha actividad, los días antes del suicidio.

Discusión

Como hemos podido observar, a pesar de la diversidad demográfica de los perfiles de los sujetos utilizados durante esta investigación, existen algunos patrones generalizados que son identificables de forma masiva (Braithwaite y col., 2016), y que nos permitirían ante una eventual creación de un instrumento más robusto poder reaccionar a tiempo ante un riesgo potencial de suicidio.

Al analizar el flujo de información provista por los mismos sujetos en sus redes sociales, se encontraron elementos que ven aumentada su frecuencia de aparición temporalmente durante momentos de crisis, episodios depresivos y demás situaciones relacionadas al suicidio. Lamentablemente, debido a las configuraciones de privacidad establecidas en la mayor parte de los perfiles de Facebook, el entrenamiento que hemos podido hacer sigue teniendo sus limitaciones. Tal vez, redes sociales con otras tendencias de privacidad como Twitter puedan arrojar datos que permitan potenciar el entrenamiento de nuestro algoritmo.

Por otro lado, es importante recalcar la trascendencia que tuvo en nuestra investigación la implementación de las redes neuronales capsulares, sobretudo a la hora de clasificar y procesar las imágenes. Comparadas con sus predecesoras las redes neuronales convolucionales, este nuevo paradigma de inteligencia artificial ha probado ser sumamente efectivo para la investigación de elementos tan complejos como los patrones de comportamiento humano. Si bien, son más lentas y complejas de calibrar, una vez que se ha llegado al punto de entrenamiento ideal, los datos se procesan de una manera más coherente y cercana a como lo hace nuestro cerebro (Iesmantas & Alzbutas, 2018).

Animamos a la comunidad a partir de nuestros

resultados y ampliar el espectro del instrumento investigativo, buscando patrones que puedan haber sido ignorados por nuestro enfoque y que finalmente se alcance un modelo más estructurado y eficiente que nos permita desarrollar mecanismos de acción temprana ante el flagelo del suicidio.

Referencias

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *CoRR*, *abs/1603.04467*. arXiv: 1603.04467. Recuperado desde <http://arxiv.org/abs/1603.04467>
- Braithwaite, S. R., Giraud-Carrier, C., West, J., Barnes, M. D. & Hanson, C. L. (2016). Validating Machine Learning Algorithms for Twitter Data Against Established Measures of Suicidality. *JMIR Mental Health*, 3(2), e21. doi:10.2196/mental.4822
- Carruthers, H. R., Morris, J., Tarrier, N. & Whorwell, P. J. (2010). The Manchester Color Wheel: development of a novel way of identifying color choice and its validation in healthy, anxious and depressed individuals. *BMC Medical Research Methodology*, 10(1), 12. doi:10.1186/1471-2288-10-12
- Iesmantas, T. & Alzbutas, R. (2018). Convolutional capsule network for classification of breast cancer histology images. *CoRR*, *abs/1804.08376*. Recuperado desde <http://arxiv.org/abs/1804.08376>
- INMLCF. (2016). Suicidios según edad. Recuperado desde <https://www.datos.gov.co/Estadisticas-Nacionales/Suicidios-seg-n-Edad-/3nc9-zxcb/data>
- IWS. (2017). Internet Usage and Population in South America. Recuperado desde <https://www.internetworldstats.com/stats15.htm>
- Jaderberg, M., Simonyan, K., Zisserman, A. & Kavukcuoglu, K. (2015). Spatial Transformer Networks. *CoRR*, *abs/1506.02025*. ar-

- Xiv: 1506.02025. Recuperado desde <http://arxiv.org/abs/1506.02025>
- Koyawala, N., Stevens, J., McBee-Strayer, S. M., Cannon, E. A. & Bridge, J. A. (2015). Sleep Problems and Suicide Attempts Among Adolescents: A Case-Control Study. *Behavioral Sleep Medicine*, 13(4), 285-295. PMID: 24654933. doi:10.1080/15402002.2014.888655
- Littlewood, D. L., Gooding, P., Kyle, S. D., Pratt, D. & Peters, S. (2016). Understanding the role of sleep in suicide risk: qualitative interview study. *BMJ Open*, 6(8). doi:10.1136/bmjopen-2016-012113
- Perlis, M. L., Grandner, M. A., Chakravorty, S., Bernert, R. A., Brown, G. K. & Thase, M. E. (2016). Suicide and sleep: Is it a bad thing to be awake when reason sleeps? *Sleep Medicine Reviews*, 29, 101-107. doi:<https://doi.org/10.1016/j.smrv.2015.10.003>
- Sabour, S., Frosst, N. & Hinton, G. E. (2017). Dynamic Routing Between Capsules. *CoRR*, *abs/1710.09829*.
- Sally, C., Margaret, W. & Holly, H. (2016). Increase in suicide in the United States, 1999–2014. *NCHS data brief*, 241.
- Sánchez-Sosa, J. C., Villarreal-González, M. E., Musitu, G. & Martínez Ferrer, B. (2010). Ideación Suicida en Adolescentes: Un Análisis Psicosocial. *Psychosocial Intervention*, 19, 279-287.
- Tubbs, A., Perlis, M., Chakravorty, S., Basner, M., Killgore, W., Gehrels, J., ... Grandner, M. (2018). Seasonal Changes in Nocturnal Suicide Risk. *Sleep*, 41, A366.
- Ursano, R. J., Heeringa, S. G., Stein, M. B., Jain, S., Raman, R., Sun, X., ... Kessler, R. C. (2015). PREVALENCE AND CORRELATES OF SUICIDAL BEHAVIOR AMONG NEW SOLDIERS IN THE U.S. ARMY: RESULTS FROM THE ARMY STUDY TO ASSESS RISK AND RESILIENCE IN SERVICEMEMBERS (ARMY STARRS). *Depression and Anxiety*, 32(1), 3-12. doi:10.1002/da.22317
- Zeiler, M. D. & Fergus, R. (2013). Stochastic Pooling for Regularization of Deep Convolutional Neural Networks. *CoRR*, *abs/1301.3557*. arXiv: 1301.3557. Recuperado desde <http://arxiv.org/abs/1301.3557>