

**Modelo de machine learning para la predicción de intentos de suicidio en la ciudad de  
Medellín**



**Estudiantes**

Santiago Otálvaro Ospina 1020492578

Santiago Escobar Casas 1214746431

**Tutor**

Javier Fernando Botia Valderrama

2508205 - Modelos de sistemas

Universidad de Antioquia

Ingeniería de Sistemas

2020-1

## **Resumen.**

El número de suicidios ha incrementado en los últimos años y probablemente va a incrementar aún más debido a las consecuencias sociales, psicológicas y económicas ocasionadas por la pandemia actual (covid-19). Por consiguiente, hemos decidido implementar un modelo de caja negra, basado en las estadísticas que hay sobre el intento de suicidio en Medellín, con el fin de realizar predicciones a futuro sobre el índice de intento de suicidio en la ciudad y generar un modelo que tenga suficiente robustez.

## **Palabras clave.**

Suicidio, Machine Learning, analítica de datos, índice de suicidio.

## **Introducción.**

El suicidio es un problema de salud pública. En estos tiempos de pandemia tan difíciles, es probable que la frecuencia con la que se cometen estos actos aumente notoriamente debido a varios factores, no solo afectando a la persona que lo intenta sino también afectando a su entorno y las personas que yacen en él. El índice de suicidio en Medellín ha venido incrementando de manera constante durante los últimos años a pesar de que en el plan municipal de salud 2016 - 2019 estaba planteado reducir la tasa de mortalidad por suicidio [1]. Esto no ha cambiado de gran manera en lo transcurrido del año 2020[2] como consecuencia de la crisis que se está viviendo en la actualidad, en muchos casos, relacionados con el aislamiento obligatorio, la violencia intrafamiliar, la falta de empleo, la crisis económica y la incertidumbre que genera la pandemia.

Se plantea implementar un modelo de caja negra (no se requiere conocimiento previo del sistema, por ende solo se requieren datos para simular el modelo) con el fin de hacer predicciones a futuro sobre el índice de suicidio en la ciudad y generar un modelo que tenga la suficiente robustez para usarse a nivel de producción.

## **Planteamiento del problema.**

Según el Boletín De Salud Mental Conducta Suicida[3], el suicidio es la muerte derivada de la utilización de cualquier método con evidencia, explícita o implícita, de que fue autoinfligida y con la intención de provocar el propio fallecimiento, se debe resaltar el hecho de que al suicidio no siempre se llega directamente sino que existen etapas o conductas previas y asociadas, tales como, la ideación suicida(conjunto de pensamientos que expresan un deseo o intencionalidad de morir), el plan

suicida(la elaboración de un plan para realizar el acto suicida, lo cual implica un alto riesgo de pasar al intento) y el intento de suicidio(conducta potencialmente lesiva autoinfligida y sin resultado fatal), todos estos conceptos se pueden englobar en lo que se define como conducta suicida.

El suicidio se considera una problemática de salud pública, por lo tanto, de ser posible se debe tratar de evitar , prestando la debida atención en cualquiera de las etapas asociadas a la conducta suicida, ya que de consumarse este hecho, puede versen afectada no solo la persona que lo intenta, sino también a su entorno; la familia, los amigos, incluso las personas que permanecen en el lugar del suceso, además de las consecuencias que le puede traer a la persona si su intento es fallido, por ejemplo: amputaciones, trastornos, traumas, rechazo, entre otras.

El gobierno incluyó esta problemática en el Plan Decenal de Salud Pública [4] 2012 -2021 en la Dimensión Social y de Salud Mental, y es uno de los eventos prioritarios para la vigilancia e intervención en salud pública en el marco de la Ley Nacional de Salud Mental [3].

Una estrategia para la prevención que se propone en este trabajo, es usar el modelo predictivo de machine learning, para que mediante la entrada de una lista con datos de interés (tupla en la cual se encuentran registrados indicios o signos de alarma) se dé la probabilidad de que un intento de suicidio pueda ocurrir para la tupla (en este caso una persona con signos de alarma) correspondiente.

## **Motivación.**

Como grupo de trabajo nuestra intencionalidad principal es crear un modelo que ayude a reducir la tasa de intento de suicidio en la ciudad de Medellín, teniendo en cuentas los factores o síntomas más relevantes e influyentes con respecto a este problema de salud pública, de manera que las personas puedan saber qué indicios o signos de alarma son más frecuentes en un caso de intento de suicidio y qué tanta probabilidad hay de que suceda, logrando así que las personas tengan una mayor prevención al momento de detectar dichos signos de alarma en alguien de su entorno. La idea surge de la necesidad de tratar de disminuir la cantidad de intentos de suicidio, ya que nos afecta a todos tanto en el aspecto social como en el aspecto psicológico, esto, teniendo en cuenta que el suicidio se considera un problema de salud publica. Adicionalmente, con el desarrollo de este proyecto pretendemos profundizar nuestros conocimientos teóricos y prácticos, sobre los modelos de caja negra y como estos pueden solventar una dificultad, o al menos contribuir a una solución.

## **Objetivos.**

- **General:** Crear un modelo supervisado de machine learning, el cual permita predecir la probabilidad de que ocurra un intento de suicidio en la ciudad de Medellín, cuando se le ingresen tuplas que contendrán los datos de interés.
  
- **Específicos:**
  - Plantear un modelo que sea cohesivo y escalable.
  - Diseñar un modelo eficiente en el manejo del volumen de datos con el cual será entrenado.
  - Entrenar el modelo, para que aumente la fiabilidad en las predicciones que este hará.
  - Verificar si el modelo de machine learning orientado al aprendizaje supervisado puede ser una estrategia útil para la prevención del intento de suicidio en la ciudad de Medellín.

## **Metodología.**

Se usará una metodología por fases donde se tendrán unas metas, las cuales, ayudarán al acercamiento del cumplimiento de los objetivos mencionados anteriormente. En cada fase se realizarán las metas concertadas y además una revisión y retroalimentación de las metas de la fase anterior.

Esta metodología por fases fue adoptada debido a la naturaleza del proyecto donde su desarrollo tiene una serie de pasos, además, la evaluación del proyecto se realiza presentado avances del mismo, por lo que la metodología escogida se adapta completamente debido a su facilidad para evaluar el cumplimiento de los objetivos.

### **1. Selección y contexto (ya realizado):**

En esta fase se pretende encontrar un problema al cual darle solución, buscar y visualizar conjuntos de datos y después de analizarlos definir si tiene la facultad de ayudarnos a alcanzar el objetivo general, comprender el escenario del problema y datos para relacionarlos adecuadamente.

### **2. Limpieza de datos:**

Luego del análisis del escenario del problema, se debe proceder a examinar detenidamente los datos obtenidos buscando lograr la reducción de dimensionalidad y la detección de datos extraños o atípicos, con el fin de mejorar la calidad del modelo desarrollado en posteriores fases, así como su eficiencia, eficacia y facilidad de uso y entrenamiento.

### **3. Selección del modelo:**

Con base al objetivo, se realizará una búsqueda bibliográfica y el estado del arte de investigaciones afines para entender distintas perspectivas y seleccionar cuáles de ellas se pueden acoger, para finalmente, decidir con cuál o cuáles proceder a realizar la simulación del comportamiento de los datos.

### **4. Entrenamiento del modelo**

En esta fase se debe separar los datos como datos de entrenamiento y datos de prueba, será necesario hacer ajustes a la importancia de cada variable buscando el modelo adecuado , las comprobaciones se harán con los datos de prueba para validar si el modelo es apropiado, o de no serlo, entrenar un nuevo modelo.

## Referencias

- [1] Betancur Pelaez, J(12 de enero de 2020). Las razones detrás del aumento de suicidios en Medellín. Recuperado de <https://www.eltiempo.com/colombia/medellin/razones-detras-del-aumento-de-suicidios-en-medellin-451046>
- [2] Instituto Nacional de Medicina Legal y Ciencias Forenses Subdirección de Servicios Forenses, 2019. Boletín Estadístico Mensual. 5th ed. [ebook] Bogotá, p.13. Recuperado de: [https://www.medicinalegal.gov.co/documents/20143/349426/Boletin\\_Mayo\\_2019.pdf](https://www.medicinalegal.gov.co/documents/20143/349426/Boletin_Mayo_2019.pdf) [Accessed 18 September 2020].
- [3] MINSALUD, Boletín de salud mental Conducta suicida Subdirección de Enfermedades No Transmisibles, 1st ed. Bogotá: Nubia Esperanza Bautista Bautista, 2020, p. 25. Recuperado de: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-conducta-suicida.pdf>
- [4] Ministerio de Salud y Protección Social and J. Acosta de la Hoz, ABC del plan Plan Decenal de Salud Pública, 1st ed. Bogotá: Fernando Chaves Valbuena, 2020, p. 82. Recuperado de: [https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/IMP\\_4feb+ABCminsalud.pdf](https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/IMP_4feb+ABCminsalud.pdf)
- [5] Instituto Nacional de Medicina Legal y Ciencias Forenses Subdirección de Servicios Forenses, 2020. Boletín Estadístico Mensual. 5th ed. [ebook] Bogotá, p.13. Recuperado de: <https://www.medicinalegal.gov.co/documents/20143/494197/5-mayo-2020.pdf> [Accessed 18 September 2020].