

# Predicción del comportamiento de una enfermedad simulada en autómatas celulares con un algoritmo propuesto en redes neuronales

May 22, 2021

## 1 Formulación del problema

Se han desarrollado gran cantidad de estudios que marcan un avance importante en el campo de la epidemiología, entre los que cabe destacar trabajos como el de sir Ronald Ross en 1911, con la incidencia y el control de la malaria, a partir de un modelo determinista, y el modelo de Kermack-McKendrick en 1927, también conocido como modelo SIR, por la relación que describe entre tres estados: susceptible-infectado-recuperado. [4]

Se han desarrollado múltiples variaciones del modelo SIR con el objetivo de analizar diferentes tipos de enfermedades, de las cuales cabe mencionar el modelo SIS, que describe la relación entre dos tipos de poblaciones: susceptibles e infectados, donde al igual que en el modelo SIR, se establecen las interacciones entre estados, a partir de un sistema de ecuaciones diferenciales acopladas.[3]

Los avances tecnológicos de las últimas décadas han sido de gran ayuda para la elaboración de modelos y simulaciones que permiten analizar características que no eran posibles con los modelos anteriormente mencionados. Además, con el reciente brote de coronavirus se identificó la necesidad de un modelo que tuviera en cuenta las relaciones sociales, lo que ocasionó que se destacarán aplicaciones de la epidemiología en redes neuronales y complejas, ambas con una fuerte influencia de la teoría de grafos. Se han realizado trabajos que analizan el impacto de la movilidad y las interacciones entre individuos usando redes neuronales, por ejemplo el descrito en [2], que toma como nodos a regiones de un país y las aristas las identifica como la movilidad humana de una región a otra, para posteriormente analizar sus implicaciones, y el desarrollado en [5] para describir la propagación del VIH en una población determinada, que tiene en cuenta la cantidad de interacciones de cada individuo en una población determinada.

La inexistencia de un algoritmo que tenga en cuenta las relaciones cercanas de cada individuo es un problema que se ha evidenciado en los avances epidemiológicos recientes, principalmente debido a la naturaleza de las bases de datos disponibles.

## 2 Justificación

A pesar del fuerte acercamiento de los modelos epidemiológicos a las dinámicas sociales con la llegada del coronavirus, se consideran interacciones a escalas que podrían ser mayores a las que se contagia la enfermedad, si bien existen modelos que analizan dinámicas de persona a persona, sus capacidades predictivas resultan limitadas por la naturaleza con la que se generan las bases de datos, como ocurre en [1, 2].

Teniendo en cuenta la aplicabilidad de los autómatas celulares y las cualidades predictivas de los modelos en redes neuronales, se pretende simular el comportamiento de una enfermedad y desarrollar un algoritmo que permita realizar pronósticos sobre su comportamiento, para responder a la pregunta: ¿Qué impacto tienen las relaciones sociales cercanas, en la propagación de una enfermedad?

Este trabajo permitirá comprender de manera más certera, como las interacciones de las personas afectan la propagación de alguna enfermedad y podrá establecer características en las bases de datos, que pueden dar lugar a análisis predictivos más realistas. Así mismo, se abre la puerta a que se propongan diferentes algoritmos con el fin de obtener diferentes resultados óptimos que alimenten el estudio epidemiológico.

## 3 Objetivos

### Objetivo general

Diseñar un algoritmo en redes neuronales que permita realizar pronósticos en el comportamiento de una enfermedad simulada con autómatas celulares, teniendo en cuenta aspectos topológicos que describan las interacciones entre individuos.

### Objetivos específicos

- Establecer la topología que describa las relaciones entre autómatas.
- Diseñar el conjunto de reglas probabilísticas que modelen las dinámicas de la enfermedad.
- Implementar las dinámicas de comportamiento de los agentes en Python, con el propósito de simular la propagación de la enfermedad.
- Identificar el conjunto de datos de entrenamiento adecuado para el algoritmo en redes neuronales.
- Definir las métricas que se van a monitorear con el algoritmo en redes neuronales.
- Diseñar e implementar un algoritmo en redes neuronales que permita realizar pronósticos para el comportamiento de la enfermedad simulada.

- Validar el comportamiento del algoritmo en ambientes topológicos distintos.

## 4 Antecedentes

## 5 Metodología

Este proyecto estará dividido en cuatro fases principales, que se articulan entre sí y con los objetivos específicos del proyecto. Las fases se ejecutarán en forma secuencial y cada una de ellas generará una serie de módulos concretos que servirán de insumo a las fases subsecuentes. A continuación, se describe brevemente cada una de las fases metodológicas junto con la articulación con los objetivos específicos del proyecto:

| Objetivos Específicos  | Fases Metodológicas  |
|--|--|
| 1. Establecer la topología que describa las relaciones entre autómatas.  | Fase 1: Definición de reglas de interacción entre individuos.      |
| 2. Diseñar el conjunto de reglas probabilísticas que modelen las dinámicas de la enfermedad.   |  |
| 3. Implementar las dinámicas de comportamiento de los agentes en Python, con el propósito de simular la propagación de la enfermedad.        | Fase 2: Simulación de la enfermedad.                               |
| 4. Identificar el conjunto de datos de entrenamiento adecuado para el algoritmo en redes neuronales.   |  |
| 5. Definir las métricas que se van a monitorear con el algoritmo en redes neuronales.  | Fase 3: Diseño e implementación del algoritmo en redes neuronales. |
| 6. Diseñar e implementar un algoritmo en redes neuronales que permita realizar pronósticos para el comportamiento de la enfermedad simulada. |  |
| 7. Validar el comportamiento del algoritmo en ambientes topológicos distintos.   | Fase 4: Entrenamiento del algoritmo y desarrollo de pruebas.       |

### Fase 1: Definición de reglas de interacción entre individuos.

La primera fase del proyecto consistirá en definir las interacciones para los diferentes participantes dentro de nuestra investigación. Partiremos de la definición de una topología definida por un sistema de vecindades locales, con el fin de simular las relaciones sociales entre individuos.

Posteriormente se hará uso de reglas totalísticas para definir las dinámicas de la enfermedad, teniendo en cuenta los comportamientos locales para cada individuo.

### Fase 2: Simulación de la enfermedad.

Una vez se establezcan las dinámicas de la enfermedad y de los individuos, se implementarán en Python usando las librerías: numpy, random, matplotlib, entre otras, para cálculos con arreglos numéricos, valores aleatorios y gráficos respectivamente.

### Fase 3: Diseño e implementación del algoritmo en redes neuronales.

Durante esta fase se definirán las métricas de interés para el estudio epidemiológico y se establecerán las neuronas que harán parte del algoritmo.

Al igual que la simulación de la enfermedad, el algoritmo definido en esta etapa, se desarrollará sobre Python usando las diferentes librerías disponibles para redes neuronales.

### Fase 4: Entrenamiento del algoritmo y desarrollo de pruebas.

Para entrenar el algoritmo propuesto, se tomarán las bases de datos simulados en la fase 2. Posteriormente se realizarán modificaciones en los parámetros de la enfermedad, el sistema de vecindades locales y la condición de frontera del conjunto de individuos, para poner a prueba las cualidades predictivas del algoritmo propuesto.

## 6 Cronograma

| Actividades a realizar  | Semanas de ejecución de cada actividad |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |
|---|--|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
|   | 1                                      | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Análisis de modelos SIS y SIR   | ■                                      |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |
| Definición e implementación del sistema de vecindades                         |  | ■ |   |   |   |   |   |   |   |    |    |    |    |    |    |    |
| Diseño e implementación de reglas de comportamiento (enfermedad e individuos) |  |   | ■ | ■ |   |   |   |   |   |    |    |    |    |    |    |    |
| Definición de métricas  |  |   |   |   | ■ | ■ |   |   |   |    |    |    |    |    |    |    |
| Diseño del algoritmo de clasificación en redes neuronales                     |  |   |   |   |   |   | ■ | ■ | ■ | ■  |    |    |    |    |    |    |
| Entrenamiento del algoritmo   |  |   |   |   |   |   |   |   |   | ■  | ■  |    |    |    |    |    |
| Diseño e implementación de los escenarios de prueba                           |  |   |   |   |   |   |   |   |   |    |    | ■  | ■  | ■  | ■  |    |
| Aplicación de pruebas sobre el algoritmo en redes neuronales y conclusiones   |  |   |   |   |   |   |   |   |   |    |    |    |    |    | ■  | ■  |
| Redacción del documento final del trabajo de grado                            |  | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■  | ■  | ■  | ■  | ■  | ■  | ■  |

## References

- [1] Cornelius Fritza, Emilio Dorigattia, and David Rugamera. “Combining Graph Neural Networks and Spatio-temporal Disease Models to Predict COVID-19 Cases in Germany”. In: *arXiv* (2021).
- [2] George Panagopoulos, Giannis Nikolentzos, and Michalis Vazirgiannis. “Transfer Graph Neural Networks for Pandemic Forecasting”. In: *arXiv* (2021).
- [3] Diego de Pereda Sebastián. “Modelización matemática de la difusión de una epidemia de peste porcina entre granjas”. MA thesis. Universidad Complutense de Madrid, 2010.
- [4] Mario Castro Ponce, Manuel de León, and Antonio Gómez Corral. “Las matemáticas contra la malaria y el modelo SIR”. In: *Matemáticas y sus fronteras* (2020).

- [5] Kao RR. “Networks and Models with Heterogeneous Population Structure in Epidemiology”. In: *Network Science*, 2010. Chap. 4.