Estimación de parámetros del modelo SIR usando redes neuronales

Andrés Camilo Fajardo Torres

Diplomado de Ciencias de datos

2022

Objetivos

Estimar los parámetros del modelo SIR para el primer pico del COVID-19 en Bogotá mediante el uso de redes neuronales.

- Programar una red neuronal que resuelva sistemas de ecuaciones diferenciales ordinarias.
- Explorar y analizar los datos de salud para el COVID-19 en Bogotá .

Table of Contents

Ecuaciones diferenciales ordinarias

Las ecuaciones diferenciales rigen el comportamiento de diversos sistemas. Estudiaremos en específico las ecuaciones diferenciales ordinarias de primer orden.

Tienen una forma funcional de la siguiente manera

$$\begin{aligned} \frac{dy}{dt} &= f(y, t) \\ y(t_0) &= y_0 \end{aligned} \tag{1}$$

donde a $y(t_0) = y_0$ se le denomina condición inicial y es lo que garantiza que la solución sea unica.

Ecuaciones diferenciales ordinarias

Entre las ecuaciones diferenciales ordinarias más empleadas tenemos

Ecuación de crecimiento logística

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) \tag{2}$$

Ley de enfriamiento de Newton

$$\frac{dT}{dt} = -k(T - T_a) \tag{3}$$

Oscilador armónico-segundo orden

$$\frac{d^2x}{dt^2} + w_0^2 x = 0 \tag{4}$$

Métodos de solución

Hay ecuaciones que se pueden solucionar de manera analítica, esto quiere decir que tienen soluciones matemáticas exactas.

Ecuación de crecimiento logística

$$N = \frac{kN_0 exp(rt)}{k + N_0(exp(rt) - 1)} \tag{5}$$

Ley de enfriamiento de Newton

$$T = T_a + (T_0 - T_a)exp(-kT) \tag{6}$$

Oscilador armonico

$$x(t) = Ae^{-iw_0 t} + Be^{iw_0 t} \tag{7}$$

Métodos numéricos

Existen métodos de aproximación que dan un resultado aproximado, con un error que depende del método y de ciertos parámetros.

Definición de derivada

- Euler
- Rugen-Kuta orden n

Redes Neuronales

Dado que las redes neuronales son aproximadores universales, se pueden emplear para la solución de ecuaciones diferenciales[2].

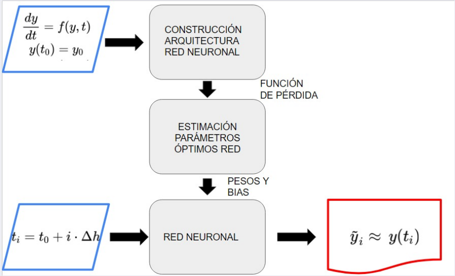

Figure: Diagrama de flujo red neuronal

Funciones de perdidas

Sea $\hat{u} = \hat{u}(W, b)$ una solución de prueba que es función del *output* de la red neuronal

$$
\begin{aligned}
j(\hat{u}) &= \sum_{i} \left(\frac{d\,\hat{u}}{dt} - f(x_I, \hat{u}) \right)^2 \\
j(\hat{\mathbf{u}}) &= \sum_{i,j} \left(\frac{d\,\hat{u}_j}{dt} - f(x_i, \hat{\mathbf{u}}) \right)^2
\end{aligned}
\tag{8}
$$

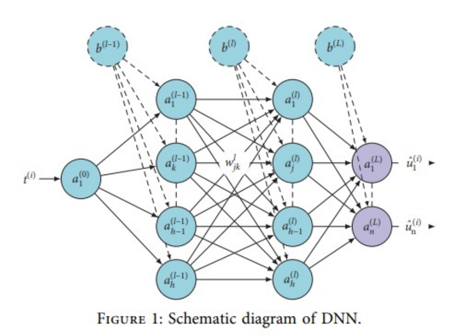

Figure: Esquema de red neuronal DNN[2]

Ecuaciones diferenciales ordinarias

Empleando una red neuronal con 2 layer ocultos, cada uno de 32 neuronas se obtiene

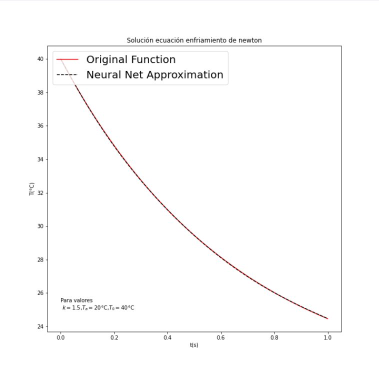

Figure: Comparación resultados usando DNN

Sistema de ecuaciones diferenciales

Un sistema de ecuaciones diferenciales son varias ecuaciones diferenciales ordinarias acopladas de la forma

$$\begin{aligned} \frac{du_i}{dt} &= f_i(u_j, t), j = 1, 2, .., n \\ u_i(t_i) &= u_{i0} \end{aligned} \tag{9}$$

Sistema de ecuaciones diferenciales

Sea el siguiente sistema de EDO asociado al modelamiento matemático de reacciones químicas irreversibles en cadena[1].

$$
\begin{aligned}
\frac{du_1}{dt} &= -k_1 u_1, \qquad & u_1(0) = 1 \\
\frac{du_2}{dt} &= k_1 u_1 - k_2 u_2, \qquad & u_2(0) = 0
\end{aligned}
\tag{10}
$$

Ecuaciones diferenciales ordinarias

Ejemplo

Usando una red neuronal de 2 hidden layer, cada uno de 32 neuronas, se obtiene para la ec. 10

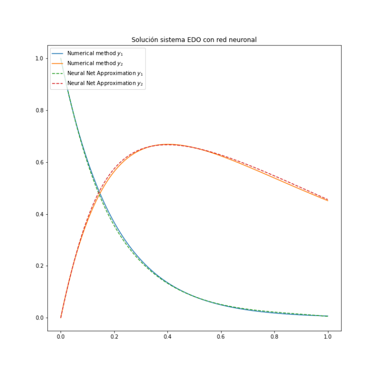

Figure: Comparación red neuronal vs métodos numericos

Estimación de parámetros

Las ecuaciones diferenciales que describen sistemas tienen parámetros que nos dan información en específico sobre el sistema de estudio .

Sea un sistema de EDO con m parámetros

$$\frac{du_i}{dt} = f_i(u_i, t, p_j), \qquad j = 1, 2, .., m \tag{11}$$

y sean unos datos observados (t_i, u_i^*) del sistema que describen las ecuaciones diferenciales.

Problema Cuál es el valor de los parámetros p que mejor ajustan los datos observados

Estimación de parámetros

Redes de neuronales Por medio de redes neuronales se puede hallar los parámetros

1 Paso Entrenar el modelo para describir los datos observados con una función de la forma

$$\hat{u}_i = u_{i0} + (t_{i0} - t)A_j^L(t_i, W, b) \tag{12}$$

con A_j el *output layer* de la red.
2 Paso A partir de 12 con parámetros W y b optimizados, se minimiza

$$j_2(p) = \sum_i \left[\frac{d\hat{u}_i}{dt}(t_i, W, b) - f_i(t_i, \hat{\mathbf{u}}, \mathbf{p}) \right] \tag{13}$$

respecto a los parámetros $\mathbf{p}$

Estimación de parámetros

Sea la ecuación 10 con parámetros desconocidos y unos datos producidos con ruido para $k = [5, 1]$

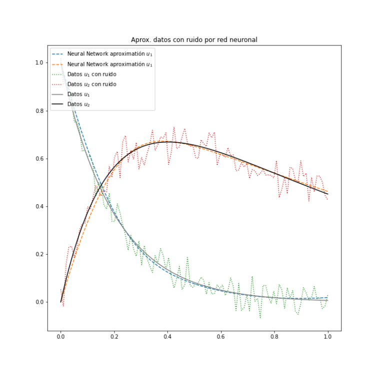

Figure: Ajuste de los datos

Estimación de parámetros

Para una red de 2 *hidden layer* cada uno de 32 neuronas, y para 1000 epochs se obtuvo los siguientes valores de k

```
EL valor de k1 es : 4.566873
EL valor de k2 es : 0.82877475
el error relativo de k1 es: 0.08662538528442383
el error relativo de k2 es: 0.17122524976730347
```

Figure: Ejemplo Estimación de parámetros usando DNN

Modelo SIR

El modelo SIR en un modelo epidemiológico que supone que la población N se mantiene constante y los sujetos obtienen inmunidad una vez contagiados.

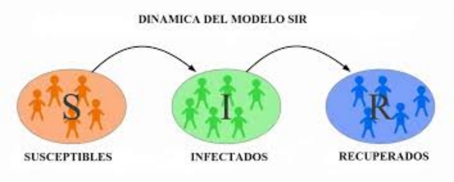

Figure: Esquema modelo SIR simple

Modelo SIR

El modelo se describe mediante las siguientes ecuaciones diferenciales

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta IS \\
\frac{dI}{dt} &= \beta SI - \gamma I \\
\frac{dR}{dt} &= \gamma S
\end{aligned}
\tag{14}
$$

con β la tasa de transmisión y γ la tasa de recuperación.

Modelo SIR

Dado que $N = S(t) + I(R) + R(t)$ y N es constante, por lo tanto $\dot{S} + \dot{I} + \dot{R} = 0$. Por lo tanto el sistema de ecuaciones diferenciales

$$\begin{aligned} \frac{dI}{dt} = \beta I(N - I - R) - \gamma I \\ \frac{dR}{dt} = \gamma S \end{aligned} \tag{15}$$

con β la tasa de transmisión y γ la tasa de recuperación.

Caso covid-19 en Bogotá

Los datos fueron obtenidos de la página del distrito: *datosabiertos.bogota.gov.co* de la fecha 14/03/2020 a 6/01/2022

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1854155 entries, 0 to 1854154
Data columns (total 11 columns):
 #   Column                      Dtype  
---  ------                      -----  
 0   CASO                        int64  
 1   FECHA_DE_INICIO_DE_SINTOMAS object 
 2   FECHA_DIAGNOSTICO           object 
 3   CIUDAD                      object 
 4   LOCALIDAD_ASIS              object 
 5   EDAD                        float64
 6   UNI_MED                     int64  
 7   SEXO                        object 
 8   FUENTE_O_TIPO_DE_CONTAGIO   object 
 9   UBICACION                   object 
 10  ESTADO                      object 
dtypes: float64(1), int64(2), object(8)
memory usage: 155.6+ MB
```

Figure: Caption

Caso covid-19 en Bogotá

Descripción de los datos

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1854155 entries, 0 to 1854154
Data columns (total 11 columns):
 #   Column                       Dtype  
---  ------                       -----  
 0   CASO                         int64  
 1   FECHA_DE_INICIO_DE_SINTOMAS  object 
 2   FECHA_DIAGNOSTICO            object 
 3   CIUDAD                       object 
 4   LOCALIDAD_ASIS               object 
 5   EDAD                         float64
 6   UNI_MED                      int64  
 7   SEXO                         object 
 8   FUENTE_O_TIPO_DE_CONTAGIO    object 
 9   UBICACION                    object 
 10  ESTADO                       object 
dtypes: float64(1), int64(2), object(8)
memory usage: 155.6+ MB
```

Figure: Caption

Descripción de los datos

Suposiciones

- El tiempo de recuperación es el mismo $t_{rec} = 45$ dias.
- El inicio de los síntomas se toma como la fecha de infección.

Análisis de los datos

Día de inicio de los sintomas

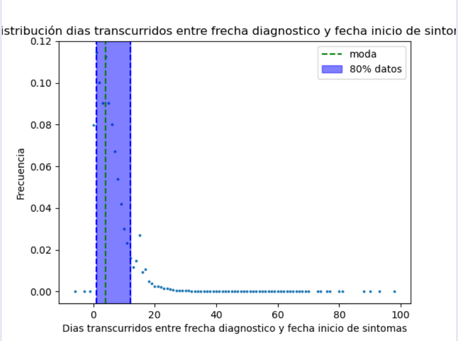

Figure: Distribución

Descripción de los datos

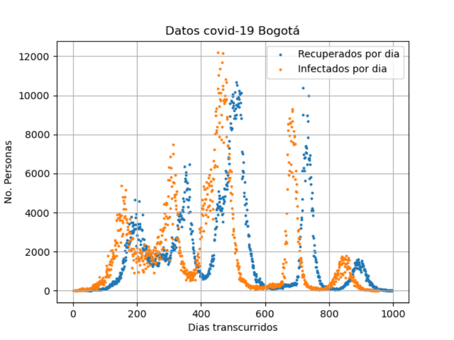

Figure: Comportamiento COVID-19 en Bogotá

Modelo SIR

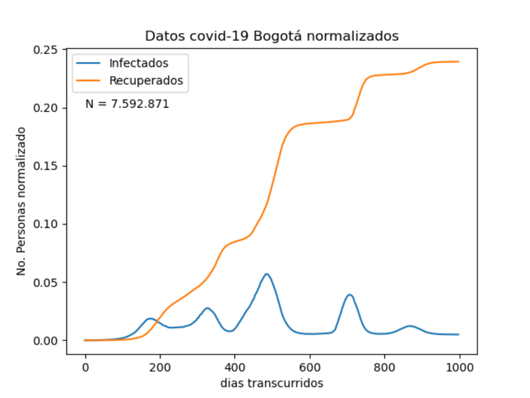

Figure: Comportamiento COVID-19 en Bogotá

Resultados

Para el primer pico, en el ajuste de los datos para 2 *hidden layer* cada uno de 62 neuronas se obtuvo

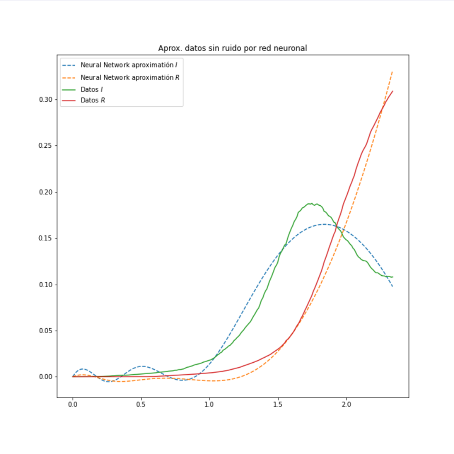

Figure: Ajuste DNN de los datos COVID-19 primer pico

Resultados

Para la estimación de los parámetros β y γ se obtienen perdidas grandes que convergen muy lento.

Soluciones

- Normalizar los datos
- N como parámetro adicional.

Bibliografia

[1] "Parameter Estimation for Dynamical Systems Using a Deep Neural Network". In: *Hindawi* (2022). DOI: https://doi.org/10.1155/2022/2014510.

[2] "Using Neural Networks to solve Ordinary Differential Equations". In: *Towards Data Science* (2022). URL: https://towardsdatascience.com/using-neural-networks-to-solve-ordinary-differential-equations-a7806de99cdd.