

# Bayesian Physics Informed Neural Networks for real-world nonlinear dynamical systems

Juan David Carrascal Ibañez

Abril 2023

# Temario

- 1 Motivación
- 2 Problema a modelar
- 3 Modelamiento vía redes neuronales
- 4 Modelamiento vía inferencia Bayesiana
- 5 Conclusiones del artículo

# Motivación

- Existe una tendencia reciente que consiste en introducir conocimiento previamente obtenido a partir de modelos físicos dentro de modelos de aprendizaje de máquina.[2]

# Motivación

- Existe una tendencia reciente que consiste en introducir conocimiento previamente obtenido a partir de modelos físicos dentro de modelos de aprendizaje de máquina.[2]
- Esto es conveniente pues el proceso de entrenamiento es menos costoso y el modelo obtenido resulta ser más robusto ante la presencia de datos con presencia de ruido o datos faltantes.[3]

# Motivación

- Además se integran técnicas de inferencia bayesiana para dar interpretación a algunos parámetros del modelo obtenido, a partir de intervalos de credibilidad.

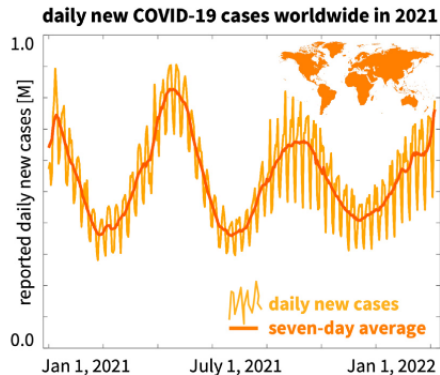
# Motivación

- Además se integran técnicas de inferencia bayesiana para dar interpretación a algunos parámetros del modelo obtenido, a partir de intervalos de credibilidad.
- Para ejemplificar el potencial de integrar estas tres técnicas, se usan datos obtenidos de la pandemia de COVID 19.

# Motivación

- Además se integran técnicas de inferencia bayesiana para dar interpretación a algunos parámetros del modelo obtenido, a partir de intervalos de credibilidad.
- Para ejemplificar el potencial de integrar estas tres técnicas, se usan datos obtenidos de la pandemia de COVID 19.
- Esto, teniendo en cuenta que las predicciones hechas con redes neuronales convencionales solían ser acertadas aproximadamente hasta una brecha de dos semanas.

# Motivación



**Fig. 1. Data - Daily new COVID-19 cases worldwide in 2021.** Thin fluctuating line illustrates the reported daily new cases of COVID-19 worldwide; thick smooth line illustrates the seven-day moving average of the daily new cases, from January 1, 2021 to December 31, 2021.



# Situación a modelar

- Partimos de una red neuronal con dos capas ocultas, tomando el tiempo  $\mathbf{t}$  como el input:

$$z_0 = t$$

$$z_1 = \sigma(W_1 z_0 + b_1)$$

$$z_2 = \sigma(W_2 z_1 + b_2)$$

$$z_3 = W_3 z_2 + b_3$$

# Situación a modelar

- Partimos de una red neuronal con dos capas ocultas, tomando el tiempo  $t$  como el input:

$$z_0 = t$$

$$z_1 = \sigma(W_1 z_0 + b_1)$$

$$z_2 = \sigma(W_2 z_1 + b_2)$$

$$z_3 = W_3 z_2 + b_3$$

- La función  $\sigma$  se tomó como  $\tanh$ .

# Situación a modelar

- El modelo físico que introduciremos está dado por la ecuación que describe a un oscilador armónico amortiguado:

$$r = \ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2x = 0$$

# Situación a modelar

- El modelo físico que introduciremos está dado por la ecuación que describe a un oscilador armónico amortiguado:

$$r = \ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2x = 0$$

- Aquí  $\zeta$  representa el radio de amortiguamiento y  $\omega_0$  la frecuencia angular. Adicionalmente consideraremos el parámetro  $\delta = c/m$ , siendo  $c$  un coeficiente de viscosidad.

# Situación a modelar

- La solución a esta ecuación viene dada por:

$$x(t) = 2A_0 \cos(\omega t + \phi) \exp(-\delta t)$$

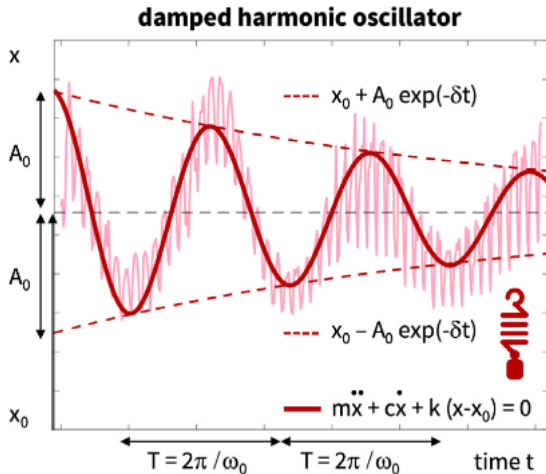
# Situación a modelar

- La solución a esta ecuación viene dada por:

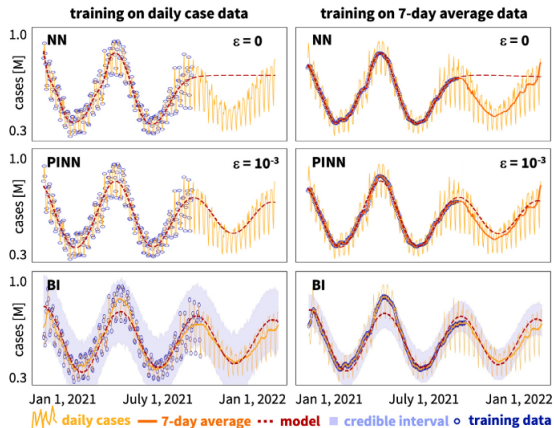
$$x(t) = 2A_0 \cos(\omega t + \phi) \exp(-\delta t)$$

- Donde,  $A_0$  es la longitud de onda,  $\omega = \sqrt{w_0^2 - \delta^2}$  es la frecuencia y  $\phi = \arctan(-\delta/\omega)$  es el angulo de fase.

# Situación a modelar



# Situación a modelar



## Comportamineto diferentes modelos



# Modelamiento vía redes neuronales

- El espacio parametral del modelo  $\Theta$ , está constituido por:

$$\Theta = \{\theta, \vartheta, \epsilon\} \quad \text{donde} \quad \theta = \{W_k, b_k\} \text{ y } \vartheta = \{c, k, x_0\}$$

# Modelamiento vía redes neuronales

- El espacio parametral del modelo  $\Theta$ , está constituido por:

$$\Theta = \{\theta, \vartheta, \epsilon\} \quad \text{donde} \quad \theta = \{W_k, b_k\} \text{ y } \vartheta = \{c, k, x_0\}$$

- Aquí  $\theta$  representa los parámetros de la red neuronal,  $\vartheta$  los parámetros del oscilador armónico y  $\epsilon$  es un peso usado para equilibrar aportación de cada uno de estos modelos.

# Modelamiento vía redes neuronales

- El espacio parametral del modelo  $\Theta$ , está constituido por:

$$\Theta = \{\theta, \vartheta, \epsilon\} \quad \text{donde} \quad \theta = \{W_k, b_k\} \text{ y } \vartheta = \{c, k, x_0\}$$

- Aquí  $\theta$  representa los parámetros de la red neuronal,  $\vartheta$  los parámetros del oscilador armónico y  $\epsilon$  es un peso usado para equilibrar aportación de cada uno de estos modelos.
- El objetivo de PINN es aprender  $\theta$  y  $\vartheta$  entrenando la red neuronal y simultáneamente resolviendo la ecuación del oscilador armónico.

# Modelamiento via redes neuronales

- Esto se logra a partir de unificar ambos problemas de optimización en uno solo mediante la función de pérdida:

$$L(\Theta; t) = (1 - \epsilon)L_{data} + \epsilon L_{phys}$$

# Modelamiento via redes neuronales

- Esto se logra a partir de unificar ambos problemas de optimización en uno solo mediante la función de pérdida:

$$L(\Theta; t) = (1 - \epsilon)L_{data} + \epsilon L_{phys}$$

- $L_{data} = \frac{1}{n_{trn}} \sum_{i=1}^{n_{trn}} ||x(t_i) - \hat{x}_i||^2$  es la función de pérdida de la red en el conjunto de entrenamiento.

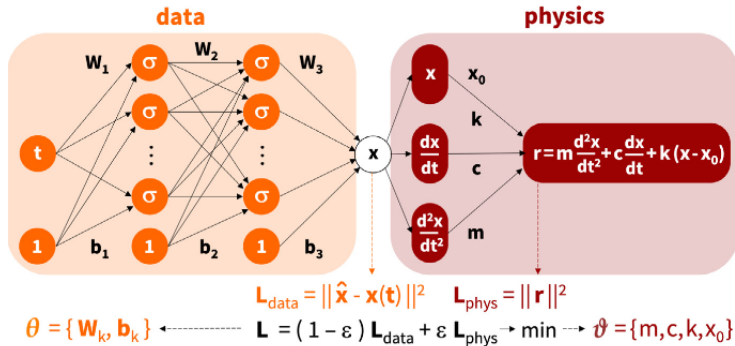
# Modelamiento via redes neuronales

- Esto se logra a partir de unificar ambos problemas de optimización en uno solo mediante la función de pérdida:

$$L(\Theta; t) = (1 - \epsilon)L_{data} + \epsilon L_{phys}$$

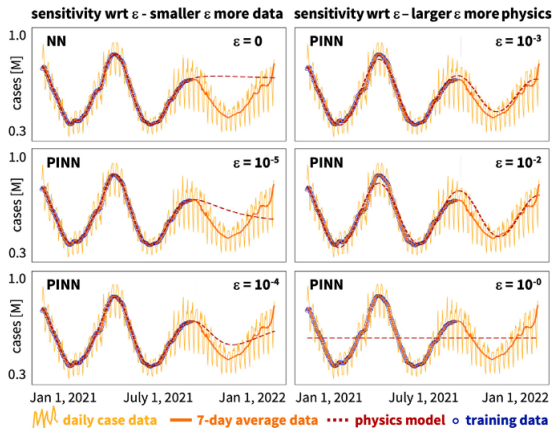
- $L_{data} = \frac{1}{n_{trn}} \sum_{i=1}^{n_{trn}} ||x(t_i) - \hat{x}_i||^2$  es la función de pérdida de la red en el conjunto de entrenamiento.
- $L_{phys} = \frac{1}{n_{smp}} \sum_{i=1}^{n_{smp}} ||\ddot{x}(t_i) + \frac{c}{m}\dot{x}(t_i) + \frac{k}{m}(x(t_i) - x_0)||^2$  es la función de pérdida usada en el oscilador armónico.

# Modelamiento vía redes neuronales



Esquema funcionamiento de una PINN.

# Modelamiento vía redes neuronales



Comparación diferentes valores de  $\varepsilon$



# Modelamiento via inferencia Bayesiana

- El objetivo de usar inferencia bayesiana es estimar la distribución posterior de un conjunto de parámetros, de manera que las estadísticas de un modelo coincidan con las del conjunto de entrenamiento, de manera que se maximice la verosimilitud de la distribución previa dada por:

$$P(\Theta|\hat{x}, r) = \frac{P(\hat{x}, r|\Theta)P(\Theta)}{P(\hat{x})P(r)} = \frac{P(\hat{x}|\Theta)P(r|\Theta)P(\Theta)}{P(\hat{x})P(r)}$$

# Modelamiento vía inferencia Bayesiana

- Aquí,  $P(\hat{x}, r|\Theta)$  es la verosimilitud del conjunto de datos  $\hat{x}$  y  $P(r|\Theta)$  es la verosimilitud de los parametros del modelo físico usado  $\mathbf{r}$ ,  $P(\Theta)$  es la distribución previa;  $P(\hat{x})$  y  $P(r)$  son las verosimilitudes marginales o también llamadas evidencias.

# Modelamiento vía inferencia Bayesiana

- Aquí,  $P(\hat{x}, r|\Theta)$  es la verosimilitud del conjunto de datos  $\hat{x}$  y  $P(r|\Theta)$  es la verosimilitud de los parametros del modelo físico usado  $\mathbf{r}$ ,  $P(\Theta)$  es la distribución previa;  $P(\hat{x})$  y  $P(r)$  son las verosimilitudes marginales o también llamadas evidencias.
- $P(\hat{x}|\Theta) = \prod_{i=0}^n p_i(\hat{x}|\Theta)$  con  $p_i(\hat{x}|\Theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\|\hat{x}-x(t_i)\|^2}{2\sigma^2}\right)$ .

# Modelamiento vía inferencia Bayesiana

- Aquí,  $P(\hat{x}, r|\Theta)$  es la verosimilitud del conjunto de datos  $\hat{x}$  y  $P(r|\Theta)$  es la verosimilitud de los parametros del modelo físico usado  $\mathbf{r}$ ,  $P(\Theta)$  es la distribución previa;  $P(\hat{x})$  y  $P(r)$  son las verosimilitudes marginales o también llamadas evidencias.
- $P(\hat{x}|\Theta) = \prod_{i=0}^n p_i(\hat{x}|\Theta)$  con  $p_i(\hat{x}|\Theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\|\hat{x}-x(t_i)\|^2}{2\sigma^2}\right)$ .
- $P(r|\Theta) = \prod_{i=0}^n p_i(r|\Theta)$  con  $p_i(r|\Theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\|r(t_i)\|^2}{2\sigma^2}\right)$ .

# Modelamiento vía inferencia Bayesiana

- Usando el mismo espacio parametral del modelo anterior  $\Theta = \{\theta, \vartheta, \epsilon\}$ , estimamos  $P(\Theta)$ , mediante:

$$P(\Theta) = P(c, k, x_0, W_k, b_k)$$

# Modelamiento vía inferencia Bayesiana

- Usando el mismo espacio parametral del modelo anterior  $\Theta = \{\theta, \vartheta, \epsilon\}$ , estimamos  $P(\Theta)$ , mediante:

$$P(\Theta) = P(c, k, x_0, W_k, b_k)$$

- Donde

$$\log(c) \sim \mathcal{N}(\mu = \log(2.2), \sigma = 0.5)$$

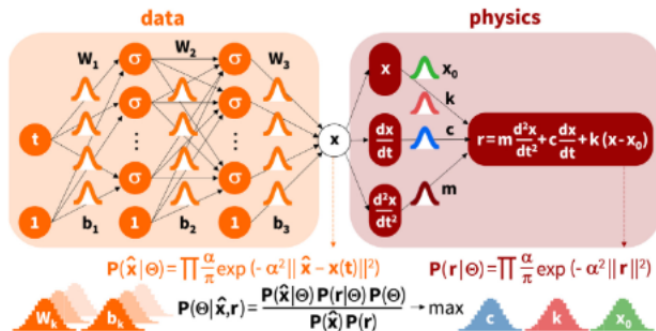
$$\log(k) \sim \mathcal{N}(\mu = \log(350), \sigma = 0.5)$$

$$\log(x_0) \sim \mathcal{N}(\mu = \log(0.56), \sigma = 0.5)$$

$$W_k \sim \mathcal{N}(\mu = 0.0, \sigma = 2.0)$$

$$b_k \sim \mathcal{N}(\mu = 0.0, \sigma = 2.0)$$

# Modelamiento vía inferencia Bayesiana



Esquema inferencia bayesiana de los parámetros de la red.

# Modelamiento vía inferencia Bayesiana

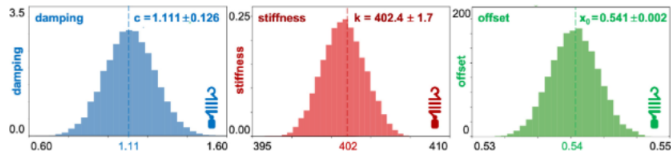


Fig. 11. Bayesian Inference. Posterior distributions of physics parameters damping  $c$ , stiffness  $k$ , and offset  $x_0$ . Inferred viscous damping coefficient of  $c = 1.111 \pm 0.126$ , stiffness of  $k = 402.4 \pm 1.7$ , and offset of  $x_0 = 0.541 \pm 0.002$ . This corresponds to a damping of  $\delta = 0.556$ , a frequency of  $\omega = 20.06$ , and a period of  $T = 0.313$  years. Simulation with training window of  $n_{\text{tr}} = 225$  days.

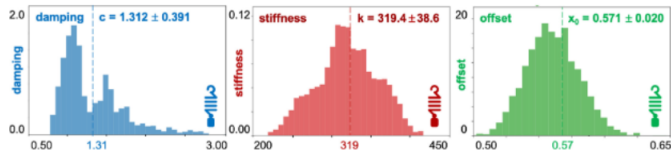
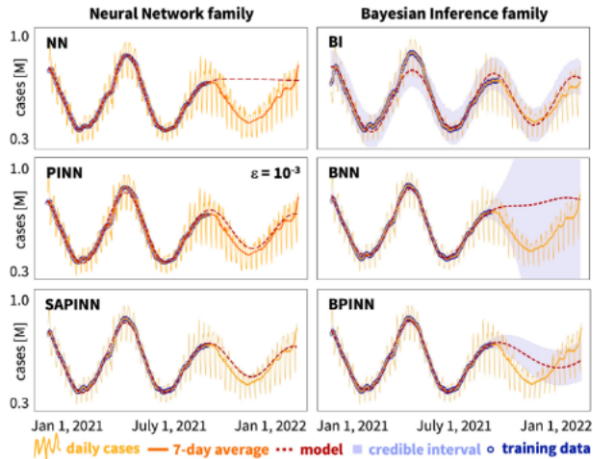


Fig. 12. Bayesian Physics Informed Neural Network. Posterior distributions of physics parameters damping  $c$ , stiffness  $k$ , and offset  $x_0$ . Inferred viscous damping coefficient of  $c = 1.312 \pm 0.391$ , stiffness of  $k = 319.4 \pm 38.6$ , and offset of  $x_0 = 0.571 \pm 0.020$ . This corresponds to a damping of  $\delta = 0.65$ , a frequency of  $\omega = 17.87$ , and a period of  $T = 0.352$  years. Simulation with training window of  $n_{\text{tr}} = 225$  days.

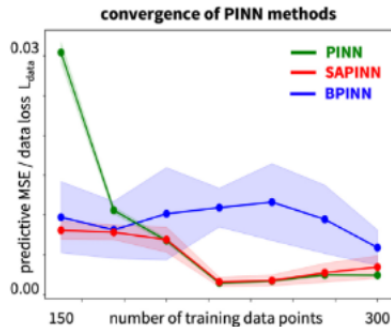


# Modelamiento vía inferencia Bayesiana



Comparación de modelos usando inferencia bayesiana.

# Conclusiones del artículo



Desempeño diferentes métodos

# Conclusiones del artículo

- Las redes neuronales por sí solas son un método simple y robusto para interpolar datos de entrenamiento pero tienen un potencial predictivo muy limitado.

# Conclusiones del artículo

- Las redes neuronales por sí solas son un método simple y robusto para interpolar datos de entrenamiento pero tienen un potencial predictivo muy limitado.
- PINN integran la información de un conjunto de datos y un modelo físico, obteniendo un buen potencial predictivo.

# Conclusiones del artículo

- Las redes neuronales por sí solas son un método simple y robusto para interpolar datos de entrenamiento pero tienen un potencial predictivo muy limitado.
- PINN integran la información de un conjunto de datos y un modelo físico, obteniendo un buen potencial predictivo.
- La inferencia bayesiana permite ajustar un modelo físico a un conjunto de datos y proporciona intervalos de credibilidad.

# Conclusiones del artículo

- Las redes neuronales por sí solas son un método simple y robusto para interpolar datos de entrenamiento pero tienen un potencial predictivo muy limitado.
- PINN integran la información de un conjunto de datos y un modelo físico, obteniendo un buen potencial predictivo.
- La inferencia bayesiana permite ajustar un modelo físico a un conjunto de datos y proporciona intervalos de credibilidad.
- BPINN proporcionan un buen ajuste y tienen buena habilidad predictiva, sin embargo son sensibles a escalamiento.

# References

- [1] Kevin Linkaa Amelie Schäferb Xuhui Mengc Zongren Zouc George Em Karniadakisc Ellen Kuhl<sup>b</sup>. *Bayesian Physics Informed Neural Networks for real-world nonlinear dynamical systems*. 2022.
- [2] Perdikaris P. Raissi M. and Karniadakis G.E. *Physics informed deep learning: Data-driven solutions of nonlinear partial differential equations*. 2017.
- [3] Yang Y. and Perdikaris P. *Adversarial uncertainty quantification in physics-informed neural networks*. pp. 136-152: J. Comput. Phys., 2019.