

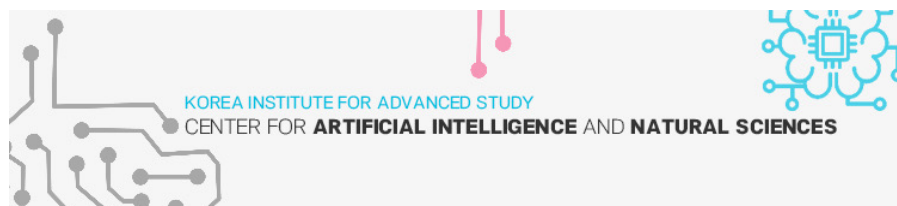
# 물리정보기반 신경망 (Physics-Informed Neural Network and Deep Ritz Method)

Jaeyong Lee

Center for AI and Natural Sciences

Korea Institute for advanced Study (KIAS)

Personal Homepage - <https://sites.google.com/view/jaeyonglee>



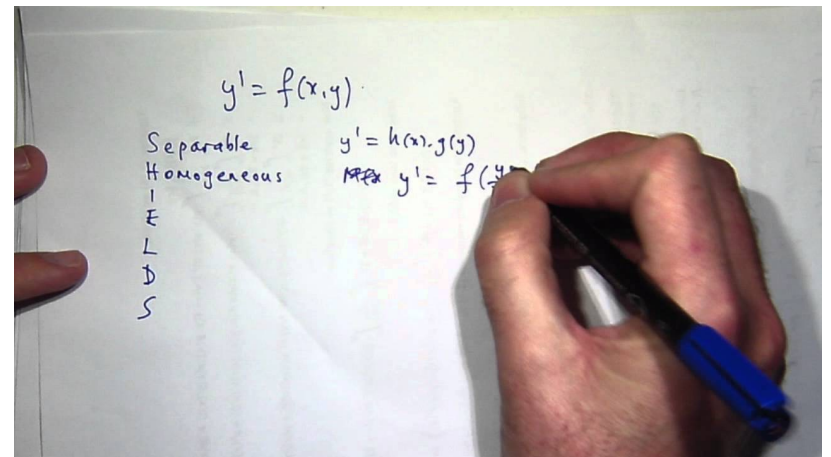
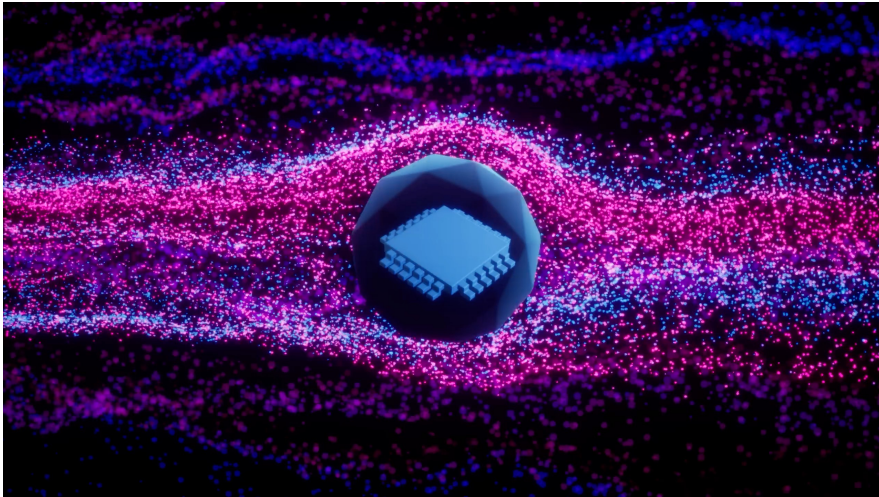
CAC 2023, June 26~27, 2023

# Contents

- **Overview**
  - Differential equations
  - Numerical methods
  - Deep neural network
- **Deep learning approach to PDEs**
  - History
  - Physics-Informed Neural Network
    - Forward Problem
    - Inverse Problem
  - Deep Ritz Method

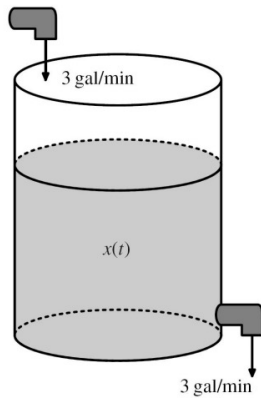
# Differential equations

- In mathematics, a differential equation is an equation that relates one or more functions and their derivatives.
  - The functions generally represent **physical quantities**.
  - The derivatives represent **their rates of change**.
  - The differential equation defines a **relationship between the above two**.



# Ordinary Differential equations (ODEs) and Partial Differential equations (PDEs)

ODE example)



Let  $x(t)$  denote the amount of salt at time  $t$ .

Balance Law:

(rate in)-(rate out)

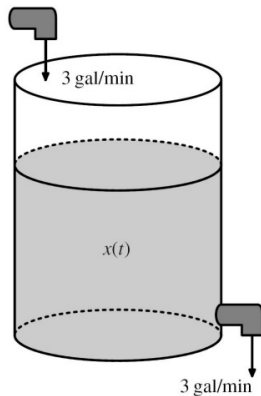
$$\frac{dx(t)}{dt} = 3 - 3x(t).$$

Initial condition (IC) :

$$x(t = 0) = 100.$$

# Ordinary Differential equations (ODEs) and Partial Differential equations (PDEs)

ODE example)



Let  $x(t)$  denote the amount of salt at time  $t$ .

Balance Law:

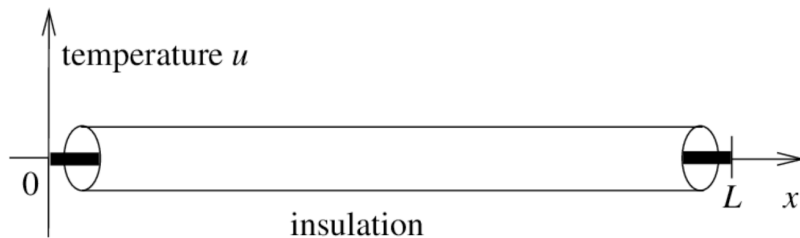
(rate in)-(rate out)

$$\frac{dx(t)}{dt} = 3 - 3x(t).$$

Initial condition (IC) :

$$x(t = 0) = 100.$$

PDE example)



Let  $u(t, x)$  denote the temperature at point  $x$  at time  $t$ .

Governing equation (heat equation) :

$$\frac{\partial u(t, x)}{\partial t} = k \frac{\partial^2 u(t, x)}{\partial x^2}.$$

Initial condition (IC) :

$$u(t = 0, x) = f(x).$$

Boundary condition (BC) :

$$u(t, x = 0) = h(t), \quad u(t, x = L) = g(t).$$

# It is Hard to solve the equation exactly!!

Let  $x(t)$  denote the amount of salt at time  $t$ .

Balance Law:

(rate in)-(rate out)

$$\frac{dx(t)}{dt} = 3 - 3x(t).$$

Initial condition (IC) :

$$x(t = 0) = 100.$$

Let  $u(t, x)$  denote the temperature at point  $x$  at time  $t$ .

Governing equation (heat equation) :

$$\frac{\partial u(t, x)}{\partial t} = k \frac{\partial^2 u(t, x)}{\partial x^2}.$$

Initial condition (IC) :

$$u(t = 0, x) = f(x).$$

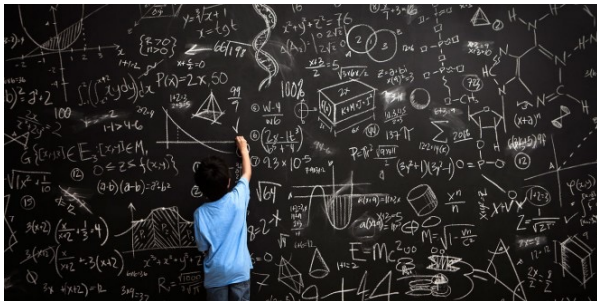
Boundary condition (BC) :

$$u(t, x = 0) = h(t), \quad u(t, x = L) = g(t).$$

# It is Hard to solve the equation exactly!!



## Numerical analysis



Let  $x(t)$  denote the amount of salt at time  $t$ .

Balance Law:

(rate in)-(rate out)

$$\frac{dx(t)}{dt} = 3 - 3x(t).$$

Initial condition (IC) :

$$x(t = 0) = 100.$$

Let  $u(t, x)$  denote the temperature at point  $x$  at time  $t$ .

Governing equation (heat equation) :

$$\frac{\partial u(t, x)}{\partial t} = k \frac{\partial^2 u(t, x)}{\partial x^2}.$$

Initial condition (IC) :

$$u(t = 0, x) = f(x).$$

Boundary condition (BC) :

$$u(t, x = 0) = h(t), \quad u(t, x = L) = g(t).$$

# Traditional numerical methods

- Finite difference method

## Finite Difference Method – Example

- Consider Poisson's equation.

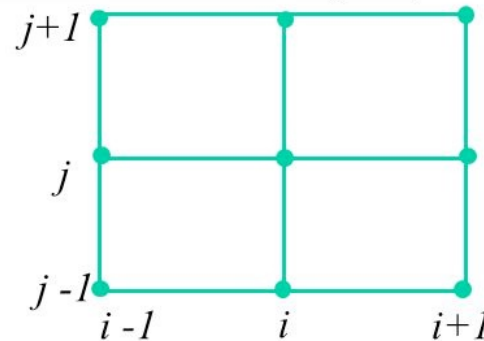
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + g = 0$$

- Discretise

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{\Delta x^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\Delta y^2} + g = 0$$

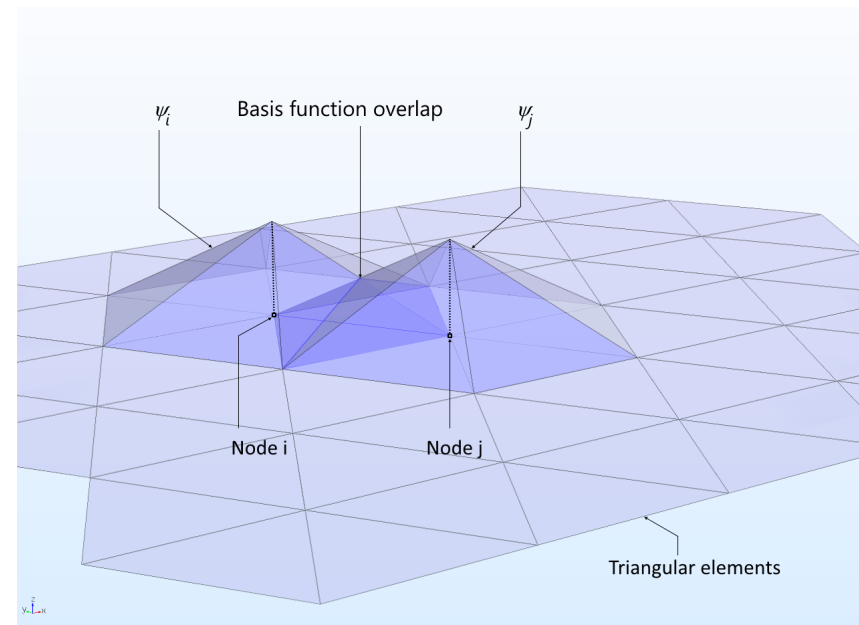
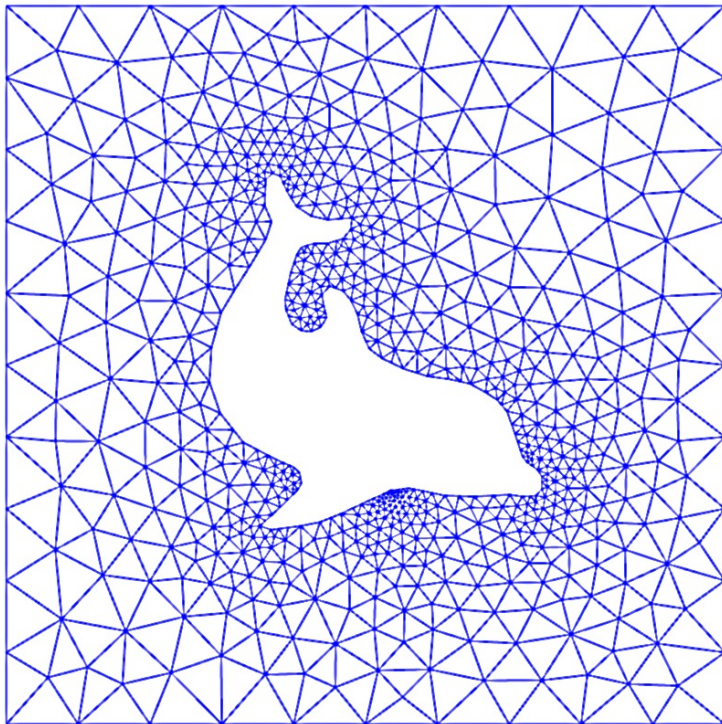
- Difference formula for each node:

$$2(\Delta x^2 + \Delta y^2)u_{i,j} = \Delta y^2(u_{i+1,j} + u_{i-1,j}) + \Delta x^2(u_{i,j+1} + u_{i,j-1}) + \Delta x^2 \Delta y^2 g$$

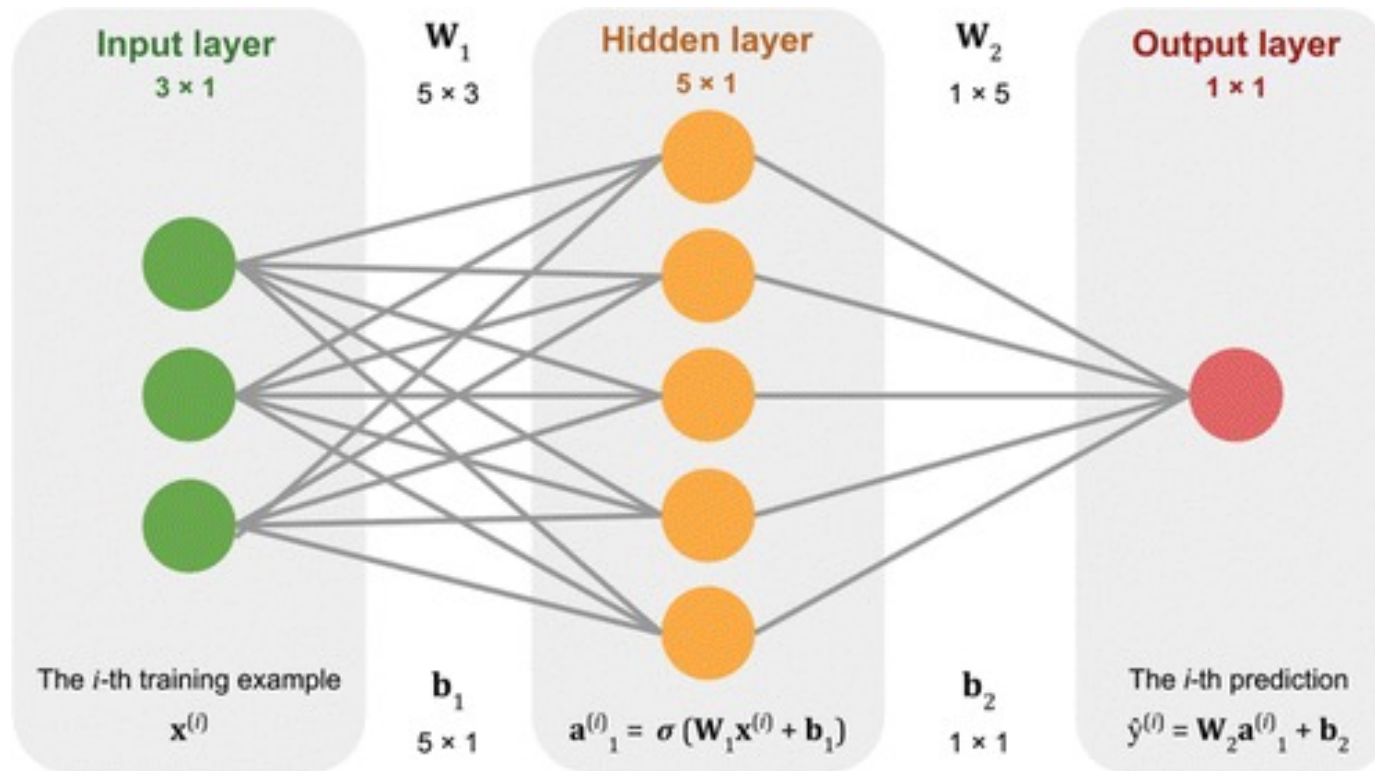


# Traditional numerical methods

- Finite difference method
- Finite element method (FEM), Finite volume method (FVM), ...



# Deep neural network



# Physics-informed machine learning???

Google Scholar

Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations

Articles

Any time

Since 2023

Since 2022

Since 2019

Custom range...

Sort by relevance

Sort by date

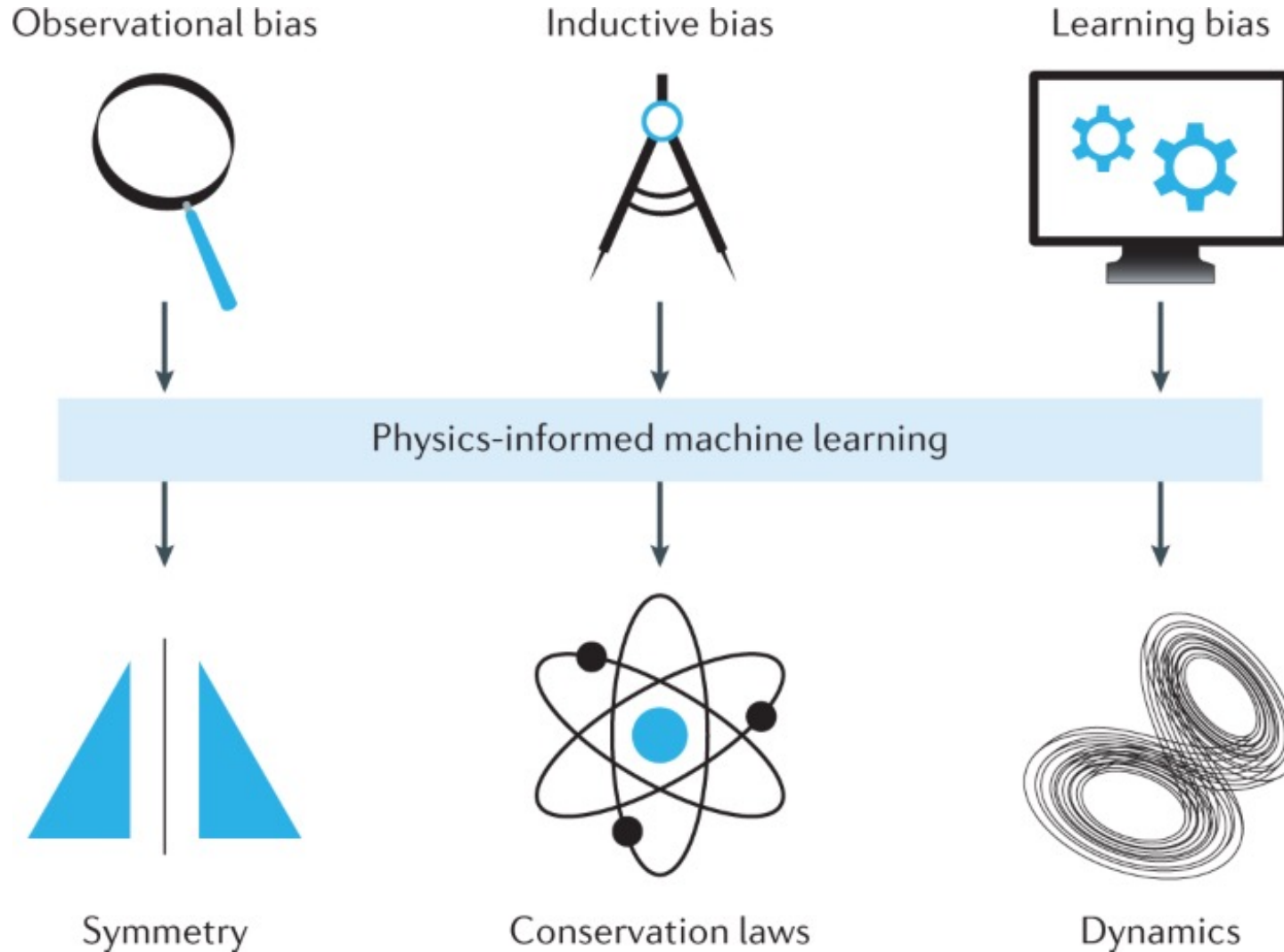
Any type

[HTML] [Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations](#)  
[M Raissi](#), [P Perdikaris](#), [GE Karniadakis](#) - *Journal of Computational physics*, 2019 - Elsevier

We introduce physics-informed neural networks—neural networks that are trained to solve supervised learning tasks while respecting any given laws of physics described by general nonlinear partial differential equations. In this work, we present our developments in the context of solving two main classes of problems: data-driven solution and data-driven discovery of partial differential equations. Depending on the nature and arrangement of the available data, we devise two distinct types of algorithms, namely continuous time and ...

☆ Save [Cite](#) Cited by 5676 [Related articles](#) [All 7 versions](#)

# Physics-informed machine learning???



# Deep learning approach to PDEs

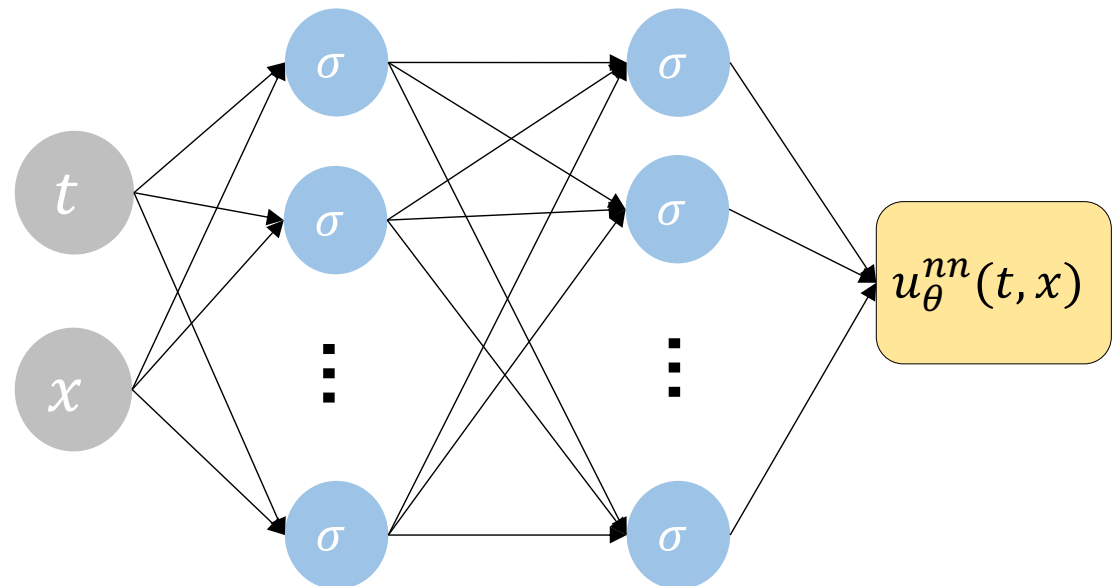
- Using neural networks directly to parametrize the solution to PDEs.
- Solve one instance of PDE at a time.
- Models
  - Physics Informed Neural Network (PINN)
  - Deep Ritz Method (DRM)

Find  $\mathbf{u}(t, \mathbf{x})$  satisfying

$$\mathcal{L}_{PDE} = f(\mathbf{u}, u_t, u_x, u_{xx}, \dots) = 0$$

$$\mathcal{L}_{IC} = u(0, \mathbf{x}) - g(\mathbf{x}) = 0$$

$$\mathcal{L}_{BC} = u|_{\partial\Omega} - h(t, \mathbf{x})|_{\partial\Omega} = 0$$



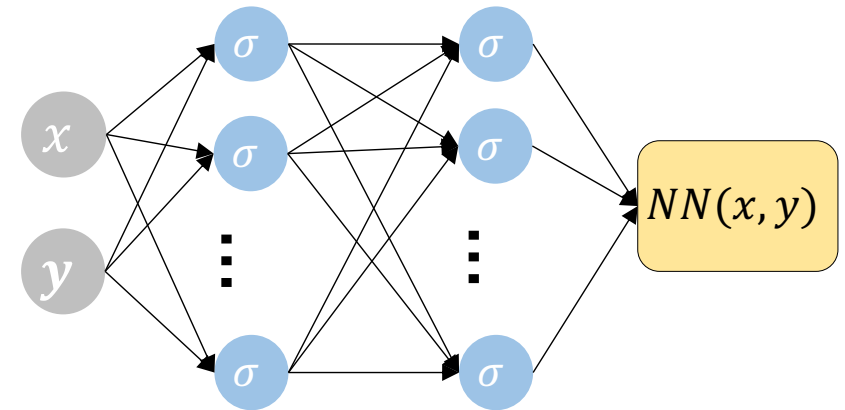
# History of PDE solver based on neural network

“Artificial neural networks for solving ordinary and partial differential equations”, IEEE Trans. Neural Netw. (1997)

Ex) Poisson equation

$$\begin{cases} \Psi_{xx} + \Psi_{yy} = f(x, y), & (x, y) \in [0, 1]^2 \\ \Psi(0, y) = f_0(y), & \Psi(1, y) = f_1(y) \\ \Psi(x, 0) = g_0(x), & \Psi(x, 1) = g_1(x) \end{cases}$$

Network Parameters :  $\vec{p}$



# History of PDE solver based on neural network

“Artificial neural networks for solving ordinary and partial differential equations”, IEEE Trans. Neural Netw. (1997)

Ex) Poisson equation

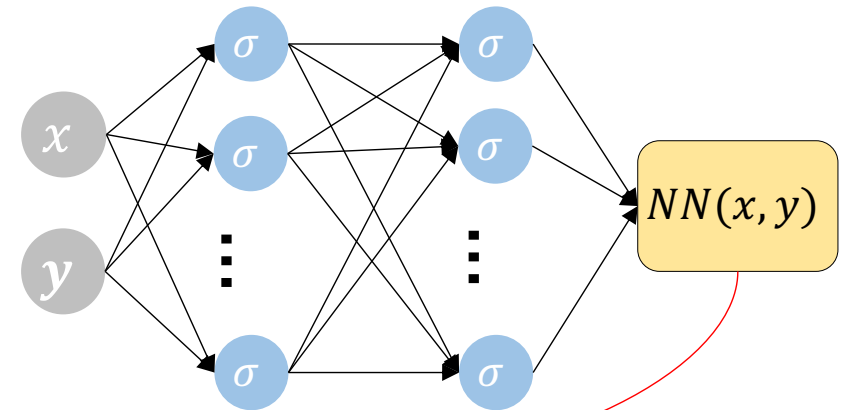
$$\begin{cases} \Psi_{xx} + \Psi_{yy} = f(x, y), & (x, y) \in [0, 1]^2 \\ \Psi(0, y) = f_0(y), & \Psi(1, y) = f_1(y) \\ \Psi(x, 0) = g_0(x), & \Psi(x, 1) = g_1(x) \end{cases}$$

They use a trial solution  $\Psi_t$  as

$$\Psi_t = A(x, y) + x(1-x)y(1-y)NN(x, y, \vec{p}),$$

where  $A(x, y)$  is chosen to satisfy the boundary conditions.

Network Parameters :  $\vec{p}$



# History of PDE solver based on neural network

“Artificial neural networks for solving ordinary and partial differential equations”, IEEE Trans. Neural Netw. (1997)

Ex) Poisson equation

$$\begin{cases} \Psi_{xx} + \Psi_{yy} = f(x, y), & (x, y) \in [0, 1]^2 \\ \Psi(0, y) = f_0(y), & \Psi(1, y) = f_1(y) \\ \Psi(x, 0) = g_0(x), & \Psi(x, 1) = g_1(x) \end{cases}$$

They use a trial solution  $\Psi_t$  as

$$\Psi_t = A(x, y) + x(1-x)y(1-y)NN(x, y, \vec{p}),$$

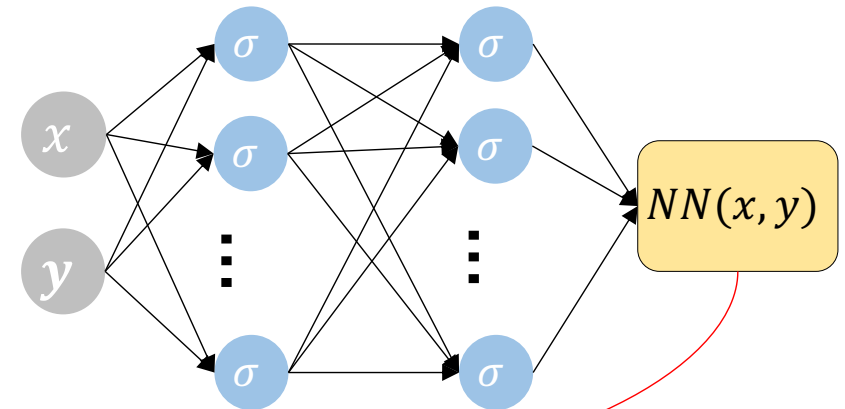
where  $A(x, y)$  is chosen to satisfy the boundary conditions.

The error to be minimized is given by

$$E(\vec{p}) = \sum_i \left\{ \frac{\partial^2 \Psi(x_i, y_i)}{\partial x^2} + \frac{\partial^2 \Psi(x_i, y_i)}{\partial y^2} - f(x_i, y_i) \right\}^2$$

How to calculate the derivatives of the output with respect to any of its inputs??  
=> Chain rule

Network Parameters :  $\vec{p}$



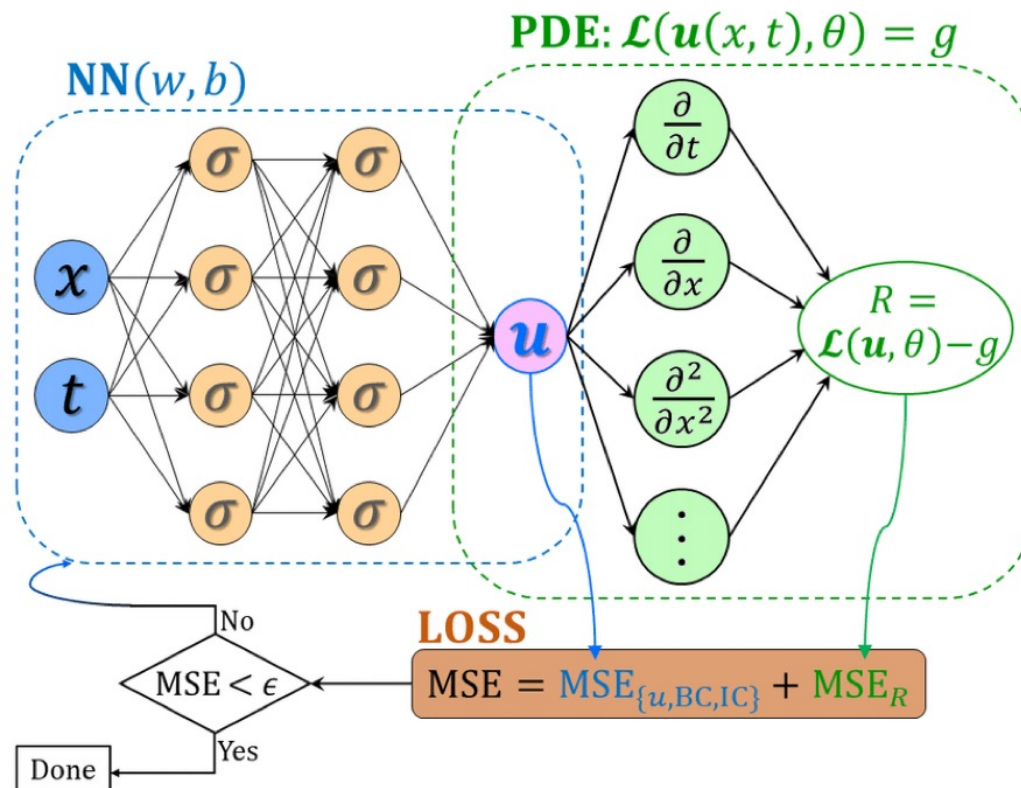
# Physics-informed neural network (PINN)

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378, 686-707.

Automatic differentiation and the back-propagation algorithm

“Here we revisit them **using modern computational tools**, and apply them to **more challenging dynamic problems** described by time-dependent nonlinear partial differential equations.”

[Schematic of a PINN ]



## Physics-informed neural network (PINN) - Forward problem

(e.g.) Burgers' equation

$$\text{Eq : } u_t + uu_x - (0.01/\pi)u_{xx} = 0$$
$$\text{IC : } u(0, x) = -\sin(\pi x) \quad x \in [-1, 1], t \in [0, 1]$$
$$\text{BC : } u(t, -1) = u(t, 1) = 0$$

# Physics-informed neural network (PINN) - Forward problem

(e.g.) Burgers' equation Eq :  $u_t + uu_x - (0.01/\pi)u_{xx} = 0$

IC :  $u(0, x) = -\sin(\pi x)$   $x \in [-1, 1], t \in [0, 1]$

**Loss function**

BC :  $u(t, -1) = u(t, 1) = 0$

$L(\theta)$

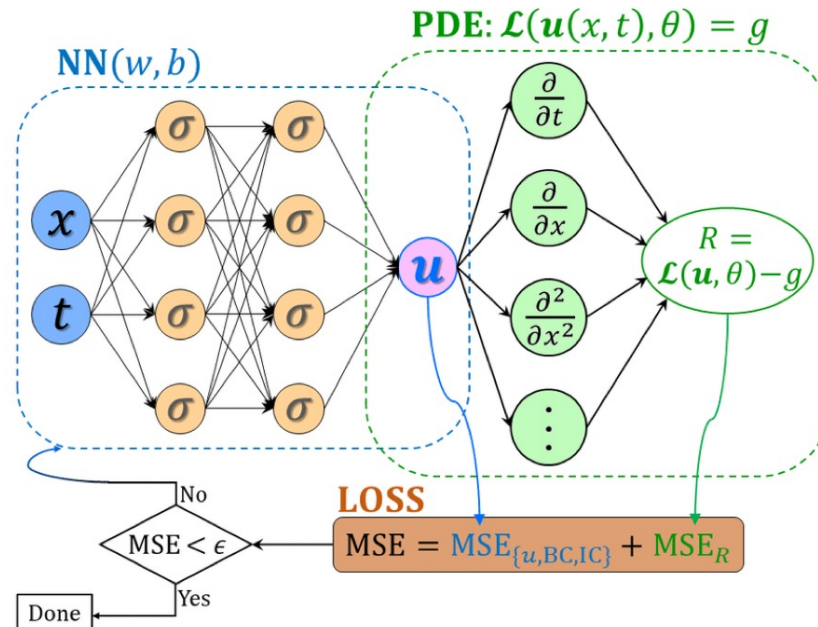
$$= \frac{1}{N_{Eq}} \sum_{i=1}^{N_{Eq}} |u_t(t^i, x^i) + u(t^i, x^i)u_x(t^i, x^i) - (0.01/\pi)u_{xx}(t^i, x^i)|^2 + \frac{1}{N_{IC}} \sum_{j=1}^{N_{IC}} |u(t^j, x^j) + \sin(\pi x^j)|^2 + \frac{1}{N_{BC}} \sum_{k=1}^{N_{BC}} |u(t^k, x^k)|^2$$

Data:

$(t_i, x_i) \in [0, 1] \times [-1, 1]$

$(t_j, x_j) \in \{0\} \times [-1, 1]$

$(t_k, x_k) \in [0, 1] \times \{-1, 1\}$



# Physics-informed neural network (PINN) - Forward problem

(e.g.) Burgers' equation Eq :  $u_t + uu_x - (0.01/\pi)u_{xx} = 0$   
 IC :  $u(0, x) = -\sin(\pi x)$   $x \in [-1, 1], t \in [0, 1]$   
 BC :  $u(t, -1) = u(t, 1) = 0$

## Loss function

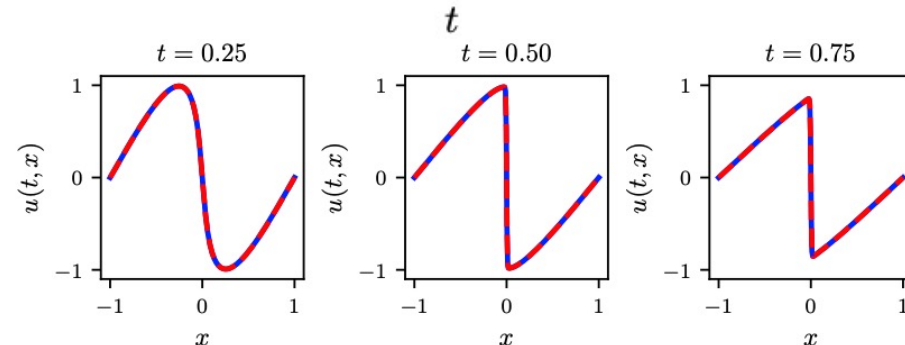
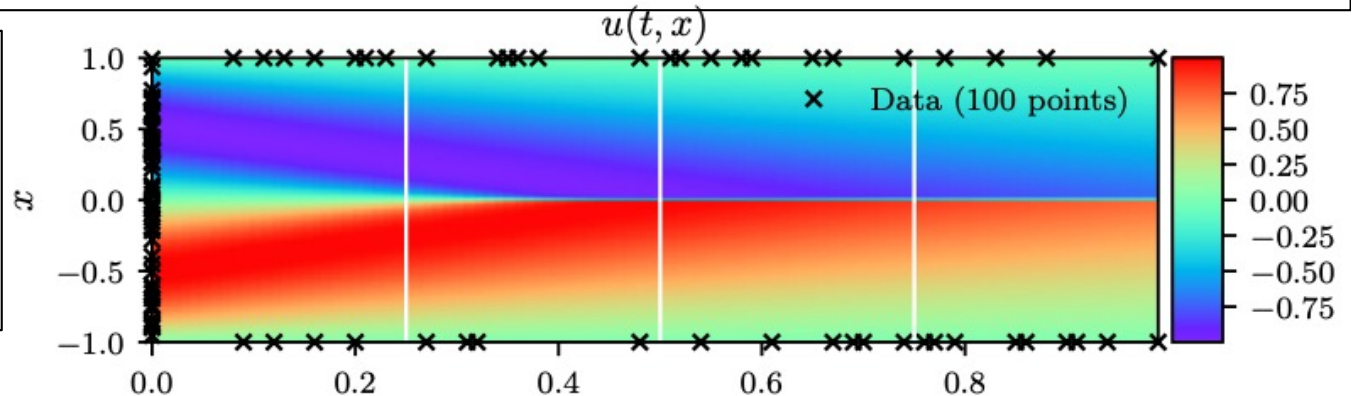
$$L(\theta) = \frac{1}{N_{Eq}} \sum_{i=1}^{N_{Eq}} |u_t(t^i, x^i) + u(t^i, x^i)u_x(t^i, x^i) - (0.01/\pi)u_{xx}(t^i, x^i)|^2 + \frac{1}{N_{IC}} \sum_{j=1}^{N_{IC}} |u(t^j, x^j) + \sin(\pi x^j)|^2 + \frac{1}{N_{BC}} \sum_{k=1}^{N_{BC}} |u(t^k, x^k)|^2$$

## Data:

$$(t_i, x_i) \in [0, 1] \times [-1, 1]$$

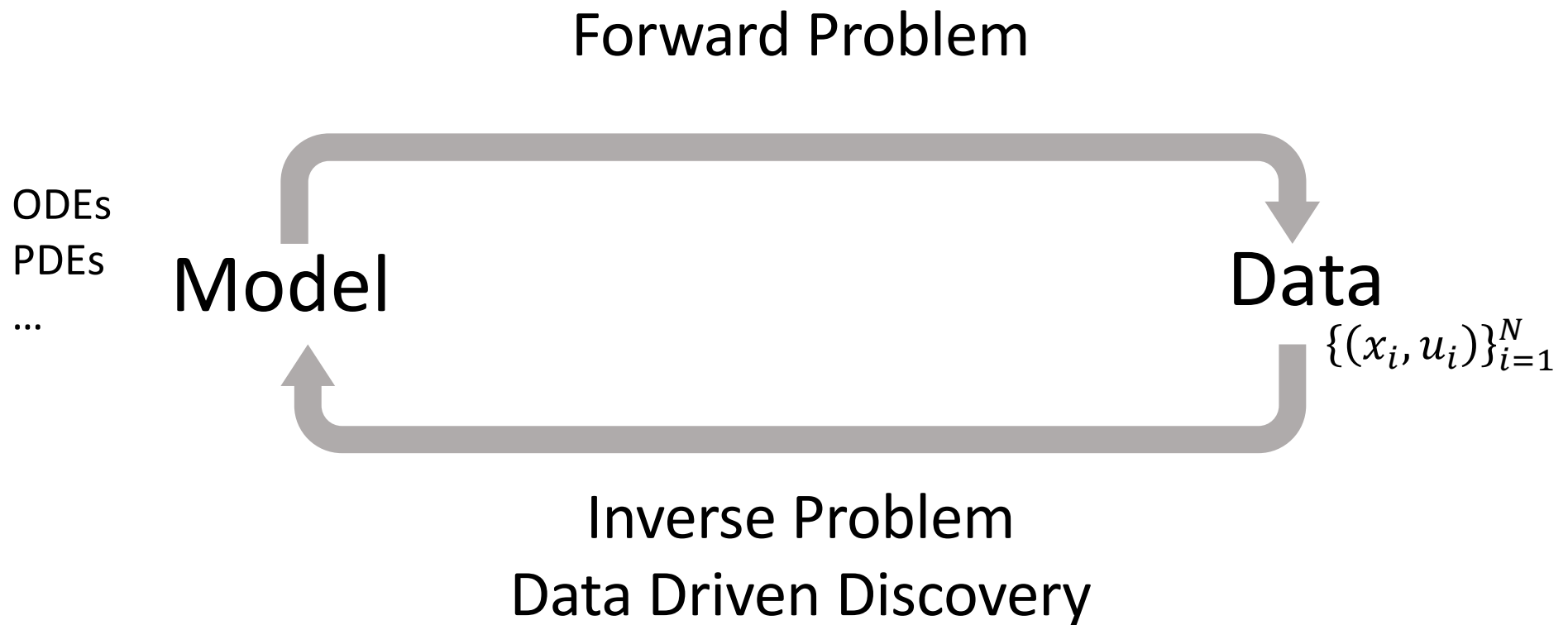
$$(t_j, x_j) \in \{0\} \times [-1, 1]$$

$$(t_k, x_k) \in [0, 1] \times \{-1, 1\}$$



— Exact — Prediction

# Physics-informed neural network (PINN) - Inverse problem???



# Physics-informed neural network (PINN) - Inverse problem

(e.g.) Burgers' equation

$$x \in [-1,1], t \in [0,1]$$

Forward problem

$$\text{Eq: } u_t + uu_x - (0.01/\pi)u_{xx} = 0$$

$$\text{IC: } u(0, x) = -\sin(\pi x)$$

$$\text{BC: } u(t, -1) = u(t, 1) = 0$$

## Physics-informed neural network (PINN) - Inverse problem

(e.g.) Burgers' equation    Eq :  $u_t + \lambda_1 u u_x - \lambda_2 u_{xx} = 0$   
IC :  $u(0, x) = -\sin(\pi x)$      $x \in [-1, 1], t \in [0, 1]$   
BC :  $u(t, -1) = u(t, 1) = 0$

  
Forward problem

$$\begin{aligned} \text{Eq : } & u_t + u u_x - (0.01/\pi) u_{xx} = 0 \\ \text{IC : } & u(0, x) = -\sin(\pi x) \\ \text{BC : } & u(t, -1) = u(t, 1) = 0 \end{aligned}$$

## Physics-informed neural network (PINN) - Inverse problem

(e.g.) Burgers' equation Eq :  $u_t + \lambda_1 u u_x - \lambda_2 u_{xx} = 0$   
IC :  $u(0, x) = -\sin(\pi x)$   $x \in [-1, 1], t \in [0, 1]$   
BC :  $u(t, -1) = u(t, 1) = 0$

### Data-driven discovery of PDE (Solving inverse problem for unknown parameter $\lambda_1$ and $\lambda_2$ )

- We want to identify unknown using an additional learning parameter.
- Assume a small amount of additional data  $\{t_l, x_l, u_l\}_{l=1}^{N_{data}}$  is given.
- Update not only **nn parameter  $\theta$**  but also the **PDE parameters  $\lambda_1$  and  $\lambda_2$** .

## Physics-informed neural network (PINN) - Inverse problem

(e.g.) Burgers' equation Eq :  $u_t + \lambda_1 u u_x - \lambda_2 u_{xx} = 0$

IC :  $u(0, x) = -\sin(\pi x)$

$x \in [-1, 1], t \in [0, 1]$

**Loss function**

BC :  $u(t, -1) = u(t, 1) = 0$

$L(\theta, \lambda_1, \lambda_2)$

$$= \frac{1}{N_{Eq}} \sum_{i=1}^{N_{Eq}} |u_t(t^i, x^i) + \lambda_1 u(t^i, x^i) u_x(t^i, x^i) - \lambda_2 u_{xx}(t^i, x^i)|^2 + \frac{1}{N_{IC}} \sum_{j=1}^{N_{IC}} |u(t^j, x^j) + \sin(\pi x^j)|^2 + \frac{1}{N_{BC}} \sum_{k=1}^{N_{BC}} |u(t^k, x^k)|^2$$

$$+ \frac{1}{N_{data}} \sum_{l=1}^{N_{data}} |u(t^l, x^l) - u_l|^2$$

Data:

$$(t_i, x_i) \in [0, 1] \times [-1, 1]$$

$$(t_j, x_j) \in \{0\} \times [-1, 1]$$

$$(t_k, x_k) \in [0, 1] \times \{-1, 1\}$$

$$(t_l, x_l, u_l) \in [0, 1] \times [-1, 1] \times \mathbb{R}$$

# Physics-informed neural network (PINN) - Inverse problem

(e.g.) Burgers' equation Eq :  $u_t + \lambda_1 u u_x - \lambda_2 u_{xx} = 0$

IC :  $u(0, x) = -\sin(\pi x)$

$x \in [-1, 1], t \in [0, 1]$

**Loss function**

BC :  $u(t, -1) = u(t, 1) = 0$

$L(\theta, \lambda_1, \lambda_2)$

$$= \frac{1}{N_{Eq}} \sum_{i=1}^{N_{Eq}} |u_t(t^i, x^i) + \lambda_1 u(t^i, x^i) u_x(t^i, x^i) - \lambda_2 u_{xx}(t^i, x^i)|^2 + \frac{1}{N_{IC}} \sum_{j=1}^{N_{IC}} |u(t^j, x^j) + \sin(\pi x^j)|^2 + \frac{1}{N_{BC}} \sum_{k=1}^{N_{BC}} |u(t^k, x^k)|^2$$

$$+ \frac{1}{N_{data}} \sum_{l=1}^{N_{data}} |u(t^l, x^l) - u_l|^2$$

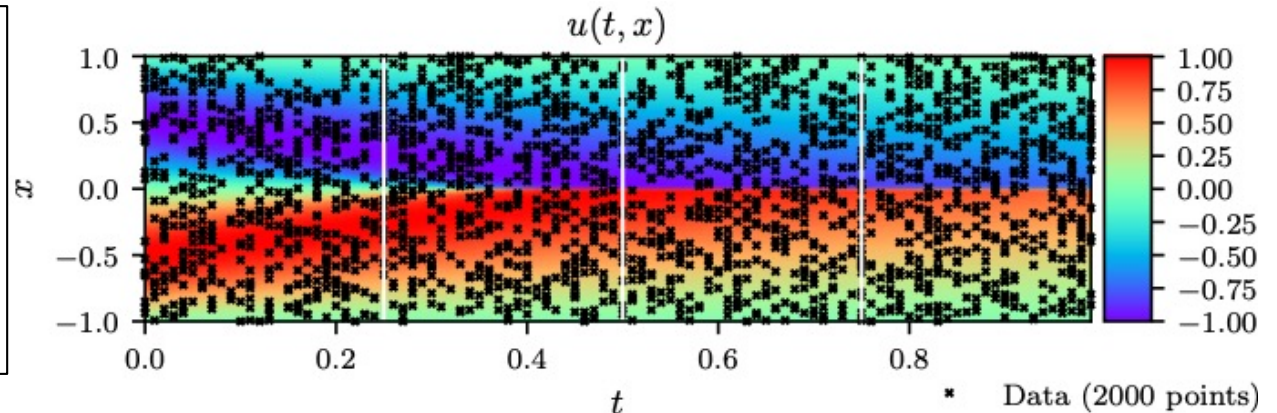
Data:

$(t_i, x_i) \in [0, 1] \times [-1, 1]$

$(t_j, x_j) \in \{0\} \times [-1, 1]$

$(t_k, x_k) \in [0, 1] \times \{-1, 1\}$

$(t_l, x_l, u_l) \in [0, 1] \times [-1, 1] \times \mathbb{R}$



# Physics-informed neural network (PINN) - Inverse problem

(e.g.) Burgers' equation Eq :  $u_t + \lambda_1 u u_x - \lambda_2 u_{xx} = 0$

IC :  $u(0, x) = -\sin(\pi x)$

$x \in [-1, 1], t \in [0, 1]$

**Loss function**

BC :  $u(t, -1) = u(t, 1) = 0$

$L(\theta, \lambda_1, \lambda_2)$

$$= \frac{1}{N_{Eq}} \sum_{i=1}^{N_{Eq}} |u_t(t^i, x^i) + \lambda_1 u(t^i, x^i) u_x(t^i, x^i) - \lambda_2 u_{xx}(t^i, x^i)|^2 + \frac{1}{N_{IC}} \sum_{j=1}^{N_{IC}} |u(t^j, x^j) + \sin(\pi x^j)|^2 + \frac{1}{N_{BC}} \sum_{k=1}^{N_{BC}} |u(t^k, x^k)|^2$$

$$+ \frac{1}{N_{data}} \sum_{l=1}^{N_{data}} |u(t^l, x^l) - u_l|^2$$

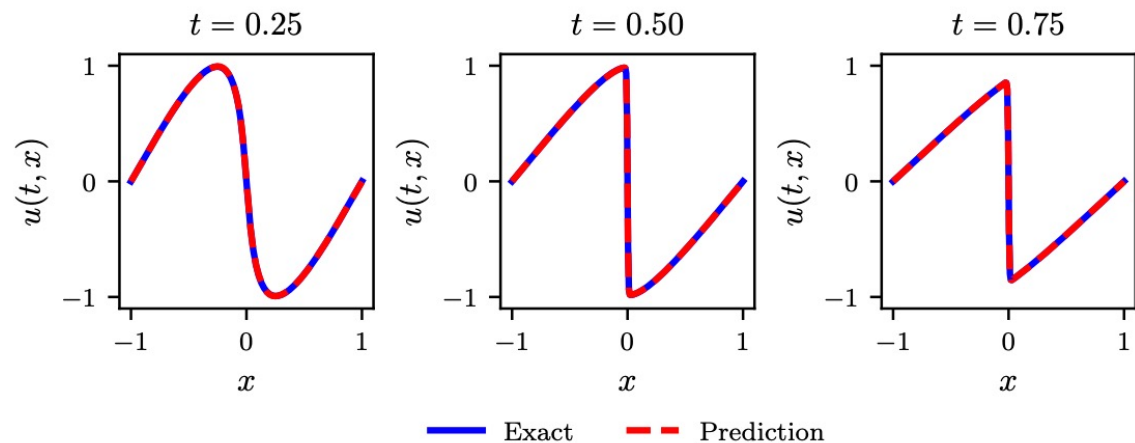
**Data:**

$(t_i, x_i) \in [0, 1] \times [-1, 1]$

$(t_j, x_j) \in \{0\} \times [-1, 1]$

$(t_k, x_k) \in [0, 1] \times \{-1, 1\}$

$(t_l, x_l, u_l) \in [0, 1] \times [-1, 1] \times \mathbb{R}$



Correct PDE	$u_t + u u_x - 0.0031831 u_{xx} = 0$
Identified PDE (clean data)	$u_t + 0.99915 u u_x - 0.0031794 u_{xx} = 0$
Identified PDE (1% noise)	$u_t + 1.00042 u u_x - 0.0032098 u_{xx} = 0$

# Deep Ritz Method (DRM)

Yu, B. (2018). The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1), 1-12.

Consider the following PDE:

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

# Deep Ritz Method (DRM)

Yu, B. (2018). The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1), 1-12.

Consider the following PDE:

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

Main idea!!

$$\text{Loss} = \min_{u|_{\partial\Omega}=0} \int_{\Omega} \left( \frac{1}{2} |\nabla u|^2 - fu \right) dx$$

# Deep Ritz Method (DRM)

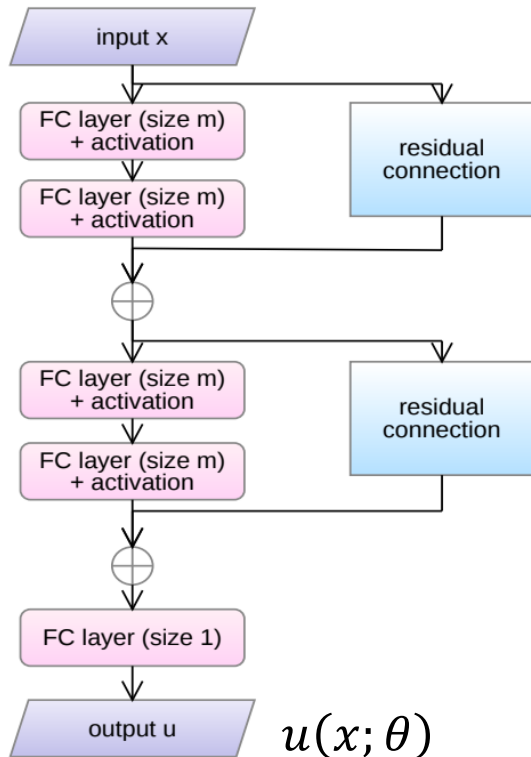
Yu, B. (2018). The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1), 1-12.

Consider the following PDE:

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

Main idea!!

$$\text{Loss} = \min_{u|_{\partial\Omega}=0} \int_{\Omega} \left( \frac{1}{2} |\nabla u|^2 - fu \right) dx$$



Deep Ritz method

$$\min_{\theta} \int_{\Omega} \left( \frac{1}{2} |\nabla u(x; \theta)|^2 - f(x)u(x; \theta) \right) dx + \beta \int_{\partial\Omega} (u(x; \theta))^2 ds$$

Penalty term for BC



# Deep Ritz Method (DRM)

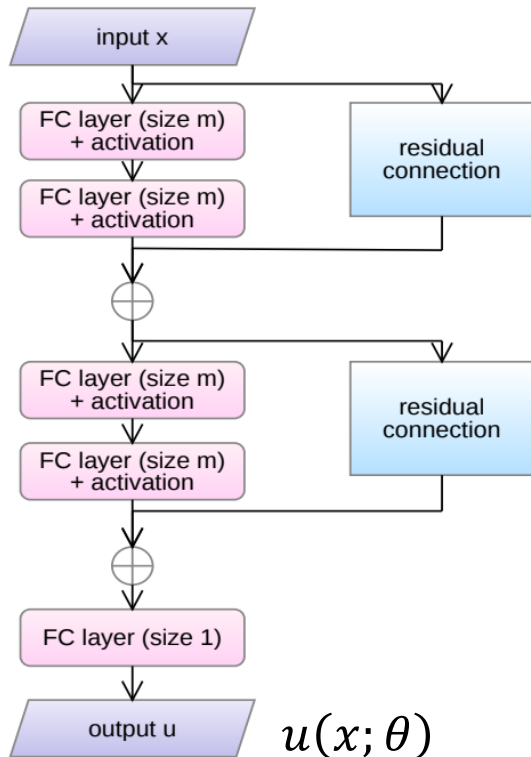
Yu, B. (2018). The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1), 1-12.

Consider the following PDE:

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

Main idea!!

$$\text{Loss} = \min_{u|_{\partial\Omega}=0} \int_{\Omega} \left( \frac{1}{2} |\nabla u|^2 - fu \right) dx$$



Deep Ritz method

$$\min_{\theta} \int_{\Omega} \left( \frac{1}{2} |\nabla u(x; \theta)|^2 - f(x)u(x; \theta) \right) dx + \beta \int_{\partial\Omega} (u(x; \theta))^2 ds$$

Penalty term for BC

Consider ( $d = 10$ )

$$\begin{aligned} -\Delta u &= 0, & x &\in (0, 1)^{10} \\ u(x) &= \sum_{k=1}^5 x_{2k-1}x_{2k}, & x &\in \partial(0, 1)^{10}. \end{aligned} \quad (15)$$

The solution of this problem is simply  $u(x) = \sum_{k=1}^5 x_{2k-1}x_{2k}$ , and we will use the exact solution to compute the error of our model later.

# Benefits and drawbacks of deep learning approach

Dockhorn, T. (2019). A discussion on solving partial differential equations using neural networks. *arXiv preprint arXiv:1904.07200*.

Table 6: Benefits and drawbacks of solving partial differential equations with neural networks.

<b>Benefits</b>	<b>Drawbacks</b>
<p><i>Expressive power:</i> Already small neural networks are able to approximate the solution of partial differential equations well.</p>	<p><i>Convergence:</i> Most probably due to optimization issues, we could not empirically show that errors decrease (with a certain convergence rate) with increasing neural network size; showing theoretical convergence seems to be even more difficult.</p>
<p><i>Ease of implementation:</i> Our algorithm is straightforward and can be easily implemented using backpropagation.</p>	<p><i>Run time:</i> At least when the BFGS optimizer is used, our simulations seem to be considerably slower than classical numerical methods. Using the proposed algorithm in combination with a first order optimization method, e.g. Adam [16], is subject to future work.</p>
<p><i>Arbitrary domains (in higher dimensions):</i> The algorithm is based on drawing random points from a domain, which can be readily extended to arbitrary domains; no triangulation of the domain is needed.</p>	<p><i>Scalability:</i> At the moment, it is unclear how well the algorithm will scale to more difficult problems, e.g., three-dimensional turbulent flows.</p>
<p><i>Freedom of approximation spaces:</i> For some classical methods and systems of PDEs, we have certain restrictions on the function spaces of the solutions, e.g., the inf-sup condition [15] needs to be satisfied for the Navier–Stokes equations when using the finite element method.</p>	<p><i>Randomness:</i> Different random initializations lead to different results of the algorithm whereas (most) classical numerical methods are deterministic. Further investigation is necessary to better exploit this non-determinism and to eventually turn it into an advantage.</p>
<p><i>Sensor data:</i> We can easily incorporate (noisy) information of sensors by adding a term to the loss function.</p>	

# Conclusion

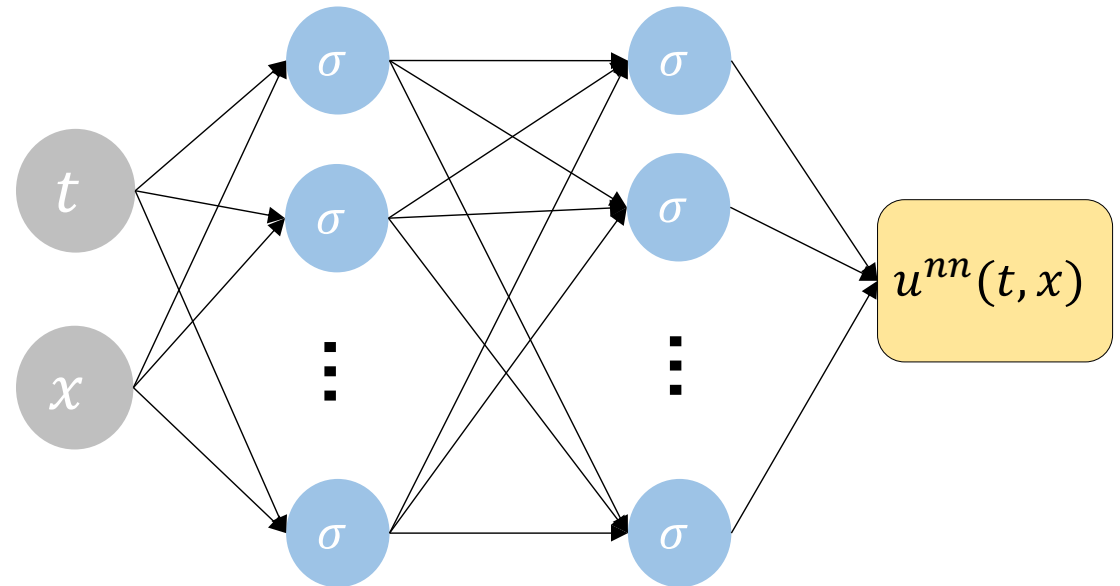
- PDE solver

Find  $\mathbf{u}(t, \mathbf{x})$  satisfying

$$\mathcal{L}_{PDE} = f(u, u_t, u_x, u_{xx}, \dots) = 0$$

$$\mathcal{L}_{IC} = u(0, \mathbf{x}) - g(\mathbf{x}) = 0$$

$$\mathcal{L}_{BC} = u|_{\partial\Omega} - h(t, \mathbf{x})|_{\partial\Omega} = 0$$



- Models

- Physics-Informed Neural Network
- Deep Ritz Method

$(\frac{1}{N} \sum)$

Let's implement

Physics-Informed Neural Network and

Deep Ritz Method!