An artificial neural network for predicting the roles of contacting bodies in computational contact mechanics

Houssam Houssein

Airthium SAS, France

Table of contents

Contact Problems

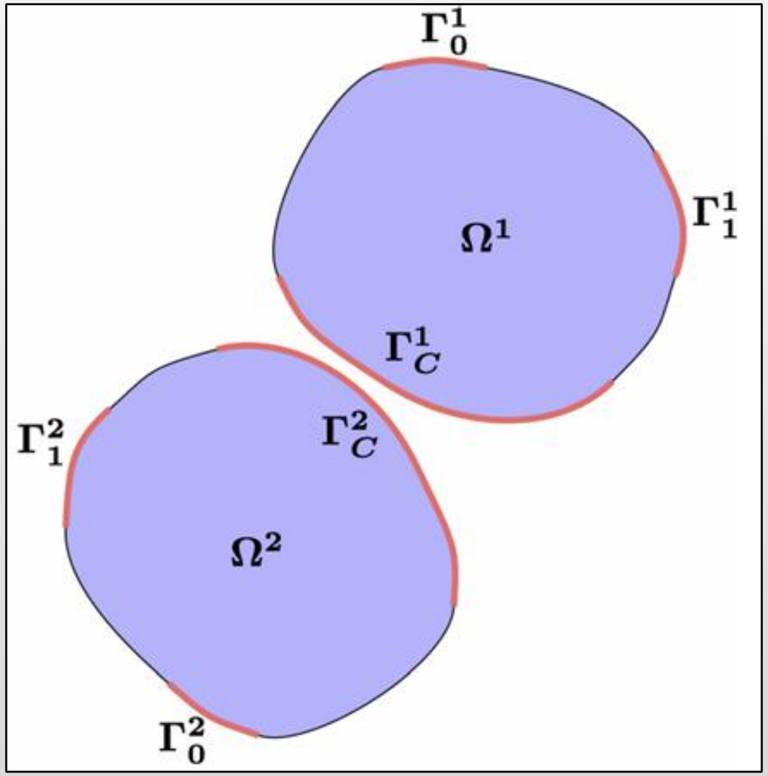
- Two bodies in contact: $\Omega^1$and Ω^2

- Balance equations for i=1,2:

$$\begin{cases} \nabla \boldsymbol{\sigma}^i + \boldsymbol{f}^i = \mathbf{0} \ \text{ in } \ \boldsymbol{\Omega}^i \\ \boldsymbol{\sigma}^i_{mn}(\boldsymbol{X}^i) = E_{mnkl} \dfrac{\partial u^i_k}{\partial X^i_l} \ \text{ in } \ \boldsymbol{\Omega}^i \\ \boldsymbol{u}^i = \mathbf{0} \ \text{ on } \boldsymbol{\Gamma}^i_0 \\ \boldsymbol{\sigma}^i \boldsymbol{n}^i = \boldsymbol{t}^i \text{ on } \ \boldsymbol{\Gamma}^i_1 \\ \boldsymbol{\sigma}^i_T = \mathbf{0} \ \text{ on } \boldsymbol{\Gamma}^1_C \cup \boldsymbol{\Gamma}^2_C \end{cases} \qquad (1)$$

- Contact conditions (Ω^1 the slave)

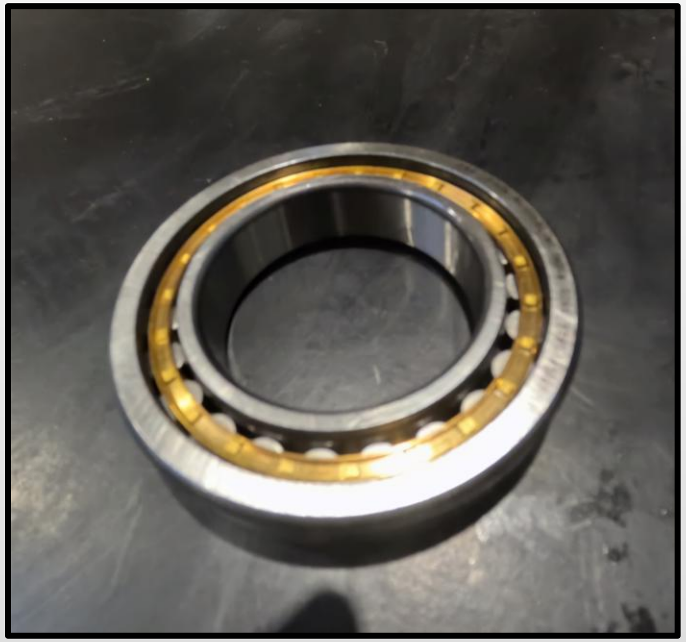
$$\begin{cases} \sigma_n := (\sigma^1 \mathbf{n}^1).\mathbf{n}^1 = \left((\boldsymbol{\sigma}^2 \circ \chi)\mathbf{n}^2\right).\mathbf{n}^2 \leq 0 \ \text{ on } \ \boldsymbol{\Gamma}^1_C \\ \left((\mathbf{u}^1 - \mathbf{u}^2 \circ \chi).\mathbf{n}^2 + g_n\right).\sigma_n = 0 \ \text{ on } \ \boldsymbol{\Gamma}^1_C \\ (\mathbf{u}^1 - \mathbf{u}^2 \circ \chi).\mathbf{n}^2 + g_n \geq 0 \ \text{ on } \ \boldsymbol{\Gamma}^1_C \end{cases} \qquad (2)$$



- $\boldsymbol{u}^i$ the displacement vector of the body Ω^i, $\boldsymbol{\sigma}^i$ the stress tensor, χ the closest point, g_n the initial gap …

Solving mechanical contact problems difficulties

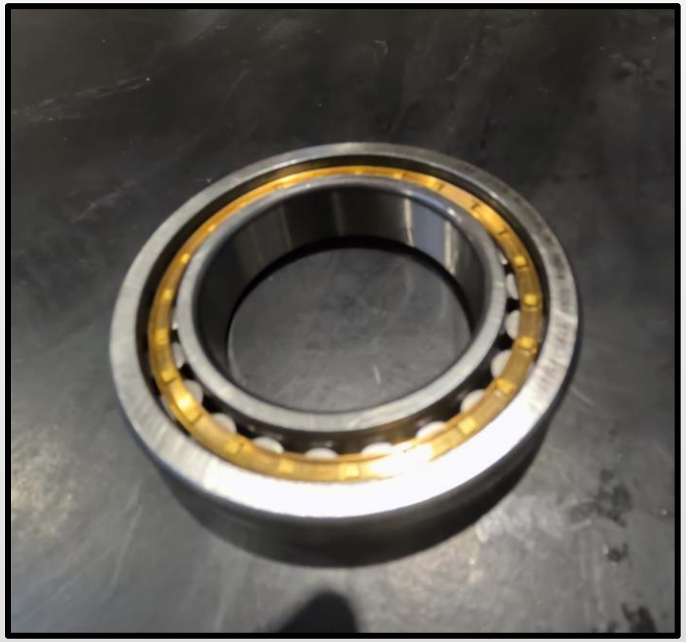
- Choose the role of each body
  - Slave or master
- Accuracy is related to the roles of the contacting bodies
- Choose the best role for each body
- How to choose a body as a slave ?
- Several recommendations exist :
  - **The body has the finest mesh**
  - **The body has the lower stiffness**
  - **The body has a curvature**
- In many situations, these recommendations do not apply
- Symmetric formulations improve accuracy but are expensive
- How to automatically assign the appropriate role–slave or master–to each contacting body ?



Cylindrical roller bearing

Solving mechanical contact problems difficulties

- Choose the role of each body
  - Slave or master
- Accuracy is related to the roles of the contacting bodies
- Choose the best role for each body
- How to choose a body as a slave ?
- Several recommendations exist :
  - **The body has the finest mesh**
  - **The body has the lower stiffness**
  - **The body has a curvature**
- In many situations, these recommendations do not apply
- Symmetric formulations improve accuracy but are expensive
- How to automatically assign the appropriate role–slave or master–to each contacting body ?
  - **Artificial neural networks (ANN)**



Cylindrical roller bearing

Artificial neural networks

- Consider a data set $\{(x_{\mathrm{m}}, y_{\mathrm{m}}) | \mathrm{m} = 1, \dots, \mathrm{M}\}$
- We want to find the relation between x_{m} and $\mathrm{y_m}$
  - Creation of an application $F_{\mathrm{w_1,b_1,\dots}}(\mathrm{x})$ which depends on several parameters $w_1, b_1, w_2, b_2, \dots$
  - Look for an optimal parameter such that
    - The search of the optimal parameters is done by a minimization process
      $\min\limits_{w_1,b_1,\dots}$ (Loss) where the function Loss measures the errors between the true value y_{m} and the function output $F_{\mathrm{w_1,b_1,\dots}}(x_{\mathrm{m}})$

- Example: Linear regression where $F_{\mathrm{w,b}}(\mathrm{x}) = \mathrm{w} \cdot \mathrm{x} + \mathrm{b}$
  - $(\mathrm{w}^*, \mathrm{b}^*) = \arg\min\limits_{\mathrm{w,b}}(\mathrm{Loss}(\mathrm{w}, \mathrm{b}))$ where

$$\mathrm{Loss}(\mathrm{w}, \mathrm{b}) = \frac{1}{\mathrm{M}} \sum_{\mathrm{m}=1}^{\mathrm{M}} \left(F_{\mathrm{w,b}}(x_m) - y_{\mathrm{m}}\right)^2 \quad (3)$$

- Artificial neural networks will represent the function F for complex relation between the data

Artificial neural networks

- Example of a three layers artificial neural network
- The output of the network is

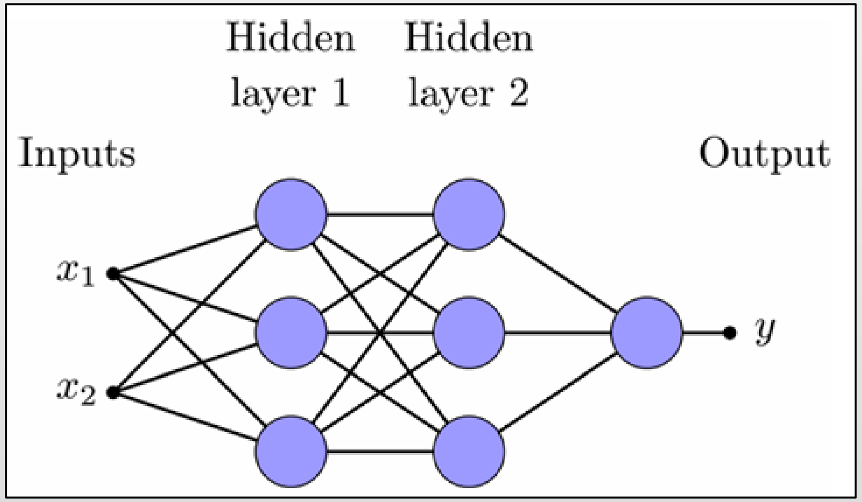
$$\mathrm{y}(\boldsymbol{x}) = F_{\boldsymbol{w}_1,\boldsymbol{b}_1,\dots}(\boldsymbol{x}) = \boldsymbol{\phi}_3(\boldsymbol{W}_3\boldsymbol{\phi}_2(\boldsymbol{W}_2\boldsymbol{\phi}_1(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1) + \boldsymbol{b}_2) + \boldsymbol{b}_3) \quad (4)$$

Weight

Activation function

Bias

Activation function Examples:
ReLU: $\phi(x) = max(x, 0)$
Log-Sigmoid : $\phi(x) = \frac{1}{1+e^{-x}}$



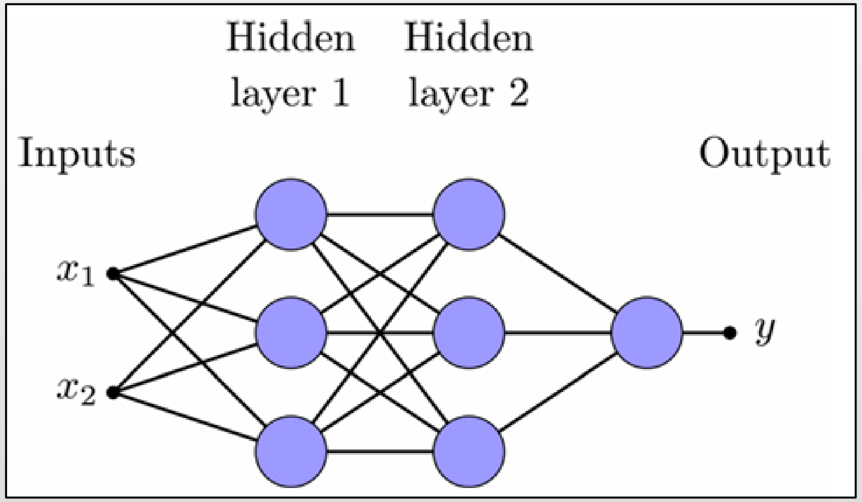
- Classification problem
  - Data set $\{(\boldsymbol{x}_{\mathrm{m}}, y_{\mathrm{m}})|\mathrm{m} = 1, \dots, \mathrm{M}\}$
  - $y_n = 1$ or 0 (First body slave ? Yes or No ?)
  - $\mathrm{y(x)} \in [0,1]$

Artificial neural networks

- Example of a three layers artificial neural network
- The output of the network is

$$\mathrm{y}(\boldsymbol{x}) = F_{\boldsymbol{w}_1,\boldsymbol{b}_1,\dots}(\boldsymbol{x}) = \boldsymbol{\phi}_3(\underbrace{\boldsymbol{W}_3}_{\text{Weight}}\boldsymbol{\phi}_2(\boldsymbol{W}_2\underbrace{\boldsymbol{\phi}_1}_{\text{Activation function}}(\boldsymbol{W}_1\boldsymbol{x} + \underbrace{\boldsymbol{b}_1}_{\text{Bias}}) + \boldsymbol{b}_2) + \boldsymbol{b}_3) \quad (4)$$

Activation function Examples:
ReLU: $\phi(x) = max(x, 0)$
Log-Sigmoid : $\phi(x) = \frac{1}{1+\mathrm{e}^{-\mathrm{x}}}$



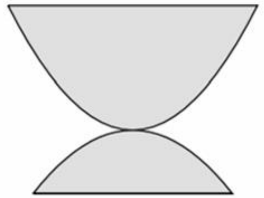
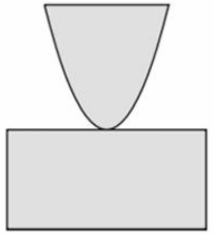
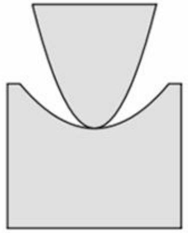
- Classification problem
  - Data set $\{(\boldsymbol{x}_\mathrm{m}, y_\mathrm{m}) | \mathrm{m} = 1, \dots, \mathrm{M}\}$
  - $y_n = 1$ or 0 (First body slave ? Yes or No ?)
  - $\mathrm{y(x)} \in [0,1]$
- The loss function is

$$Loss(\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3) = -\frac{1}{\mathrm{N}}\sum_{\mathrm{m}=1}^{\mathrm{M}}(y_\mathrm{m}\log(\mathrm{y}(\boldsymbol{x}_{\boldsymbol{n}})) + (1 - y_\mathrm{m})(\log(1 - \mathrm{y}(\boldsymbol{x}_\mathrm{m}))) \quad (5)$$

- Minimization of the loss function to determine $\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ => **Training**
- $Accuracy = \frac{n_D}{\mathrm{M}} \times 100$ where n_D is the number of data correctly classified

Domain of the study

- 2D contact problems subjected to small deformations
- No friction
- Relation between the properties of the contacting bodies and the choice of slave and master for each body. More precisely, can we take the first body as the slave? Yes or no?
- Three contact cases

Case 1	Convex structures	
Case 2	A convex structure and a rectangular block	
Case 3	Convex and concave structures	

Domain of the study

- Properties ratios:

$$\begin{cases} r_n = \dfrac{\frac{n_C^1}{S_1}}{\frac{n_C^2}{S_2}} \\ r_{curv}^1 = \dfrac{R}{R_{curv}^1} \text{ with } R = 1 \text{ mm} \\ r_{curv}^2 = \dfrac{R}{R_{curv}^2} \\ r_{Dp} = \dfrac{D_p^1}{D_p^2} \\ r_E = \dfrac{E_1}{E_2} \end{cases} \quad (6)$$

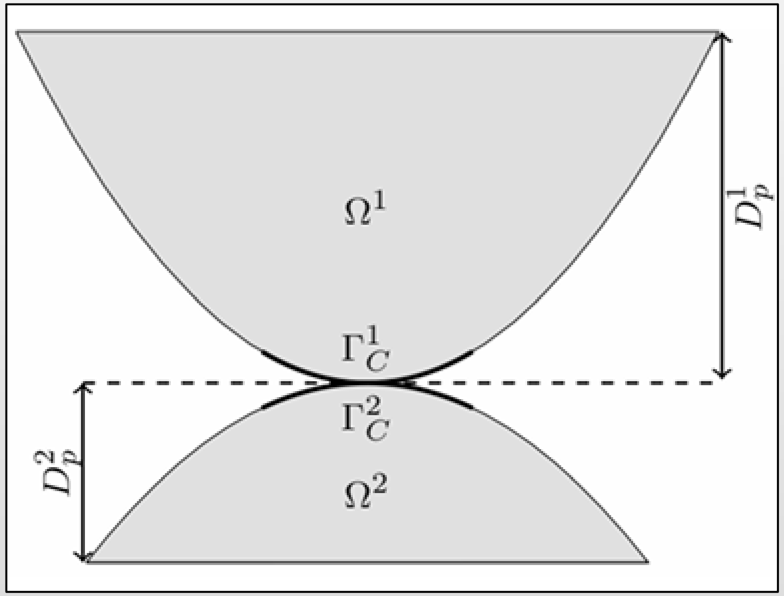
n_C^1 → Number of nodes on the potential contact surface $\boldsymbol{\Gamma}_C^1$

S_1 → Area of the potential contact surface $\boldsymbol{\Gamma}_C^1$

R_{curv}^2 → Radius of curvature of the potential contact surface $\boldsymbol{\Gamma}_C^2$

D_p^2 → Depth of the body 2

E_1 → Young's modulus of the body 1



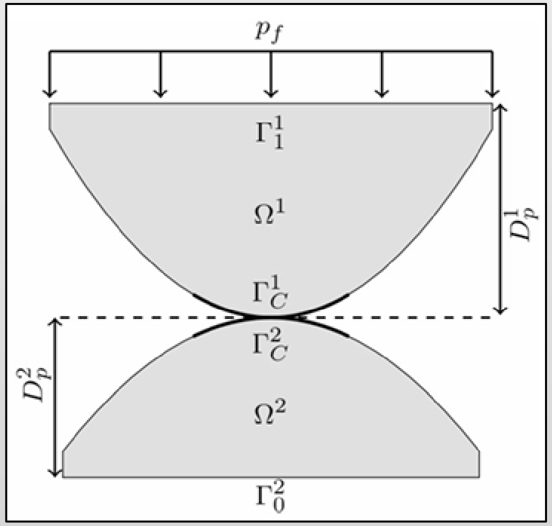
- Look for an application $F : \mathbb{R}^5 \to \{0,1\}$ such that

$$F(r_n, r_{curv}^1, r_{curv}^2, r_{Dp}, r_E) = \begin{cases} 1 \text{ if the body 1 can be taken as slave} \\ 0 \text{ otherwise} \end{cases}$$

- F will be the artificial neural network to determine

Data generation

- Hertz's assumptions (Johnson, 1987) :
  - Small strains and frictionless contact are considered
  - The contact surfaces are smooth enough
  - The dimension of the contact surface is small compared to the dimension of each body
  - The dimension of the contact surface is small compared to the radius of curvature of each body's contact surface.
- Random properties for each case
  - Meshes, Young's modulus, slopes of the two contact surfaces, pressure, ...
- For each random configuration (sample)
  - The theoretical contact pressure can be computed using Hertz theory
  - Two contact simulations are considered
    - First body as slave
    - Second body as slave
  - The results of each of the two contact simulations are compared with the theoretical results (errors err_1 and err_2)
  - Decision about the contact role of the first body (slave or not)
  - $\mathrm{r_n}, \mathrm{r^1_{curv}}, \mathrm{r^2_{curv}}, \mathrm{r_{Dp}}, \mathrm{r_E}$, 1 or $\mathrm{r_n}, \mathrm{r^1_{curv}}, \mathrm{r^2_{curv}}, \mathrm{r_{Dp}}, \mathrm{r_E}$, 0 is stored



Convex bodies

Data generation

- Some data are discarded (when errors err_1 and err_2 are relatively equal)
- All data violating Hertz's assumptions are also discarded.
- NTS (node to segment) discretization where the nodes of the slave body are prevented to penetrate is employed
- The contact algorithm is a constrained minimization one, where the total potential energy is minimized subject to non-penetration constraints, and the interior point method is used to solve this problem (Houssein, 2022)
- Data generation is done using FreeFEM (Hecht, 2012)
- Details can be found in the paper (Houssein, 2025)

Neural network architecture

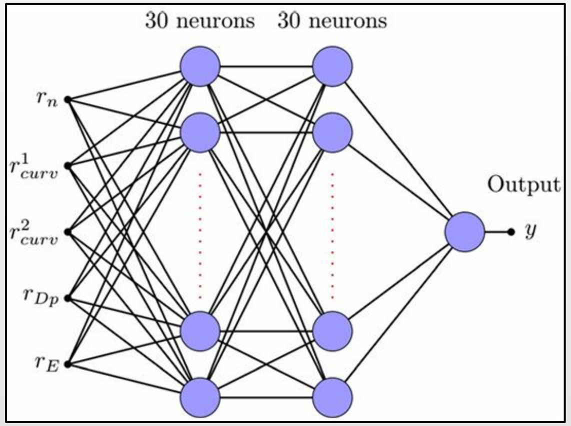
- Construction of a neural network for each contact case with five inputs corresponding to $\mathrm{r_n}, \mathrm{r^1_{curv}}, \mathrm{r^2_{curv}}, \mathrm{r_{Dp}}, \mathrm{r_E}$, and with only one output $= \begin{cases} 1 \text{ if the body 1 can be taken as slave} \\ 0 \text{ otherwise} \end{cases}$
- Preprocessing steps on the data (logarithm)
- Generated data:
  - 70% : Training data, 15% : Validation data, 15% : Test data
- The activation function of the final layer is set to be a sigmoid function, which gives values between 0 and 1, on the other hand, the remaining activation functions in the network are the ReLU functions
- Adam algorithm (Kingma, 2014) with a stepsize $\alpha = 0.001$ and hyper-parameters $\beta_1 = 0.9, \beta_2 = 0.999$ was employed to train the three neural networks
- For each contact case:
  - Different candidate neural networks with varying architectures
    - The number of training epochs (i.e., iterations in the minimization process) was chosen to maximize accuracy on the validation set and prevent overfitting.
  - The architecture achieving the highest accuracy on the training, validation, and test sets was selected as the best-performing model.

Neural network architecture

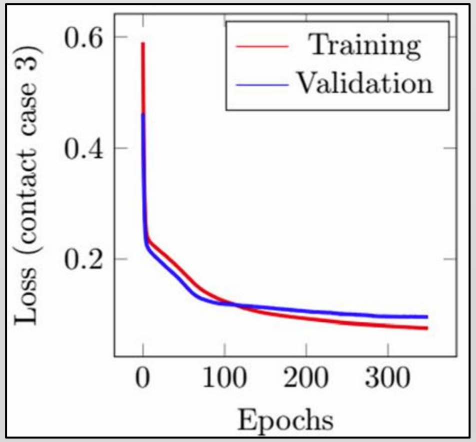
- Keras (Chollet et al., 2015) and TensorFlow (Abadi et al., 2015)were the employed tools for the creation of our neural networks.
- Optimal ANN for each case

Contact case	Network	Training accuracy(%)	Validation accuracy(%)	Test accuracy(%)	Epochs number
1	(5,30,30,1)	98.28	96.90	97.80	400
2	(5,30,1)	98.07	98.20	99.00	200
3	(5,30,30,1)	96.93	96.80	96.80	350

- ANN for contact case 3 can be employed for contact case 2
- Inputs importance (some inputs can be omitted)



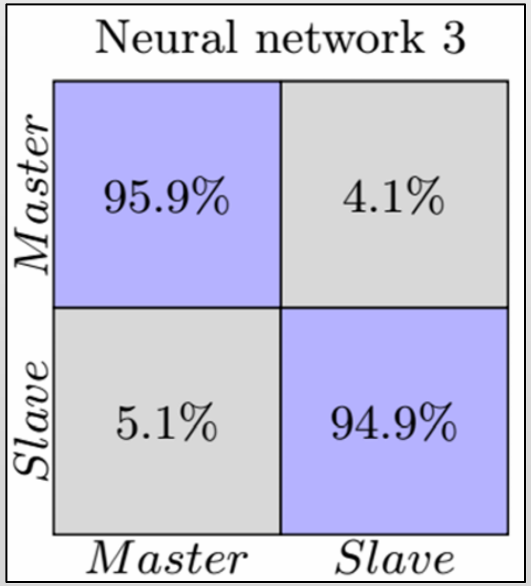
ANN architecture for contact cases 1 and 3



Loss curve vs Epochs number

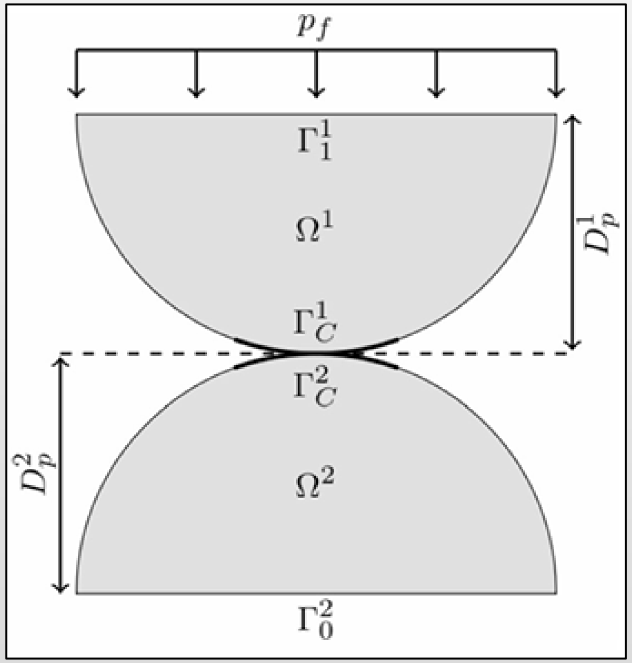
Neural network architecture

- Confusion matrix is a useful tool for binary classification problems
- Confusion matrix of the third ANN for the test data of contact cases 3
  - 95.9 % of the master data were classified correctly
  - 94.9 % of the slave data were classified correctly
  - 5.1 % of slave data were misclassified as master
  - 4.1 % of master data were misclassified as slave

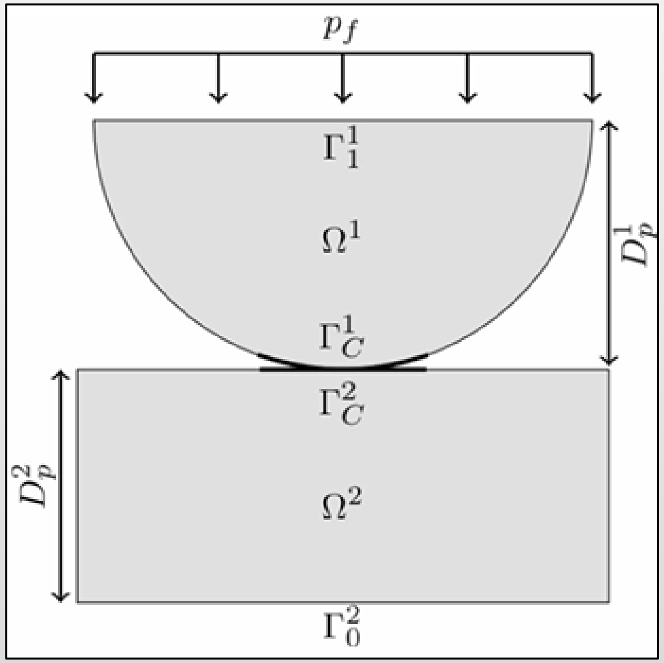


Confusion matrix of the third ANN for the test data

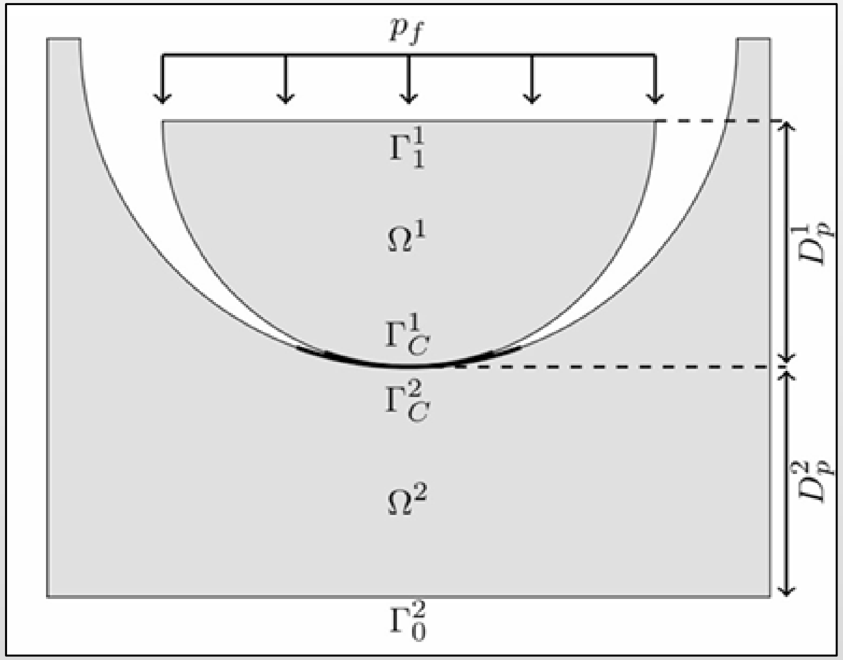
Additional validation



- **Contact between two cylinders**
- **500 samples**
- **Accuracy $= 96.40\%$**



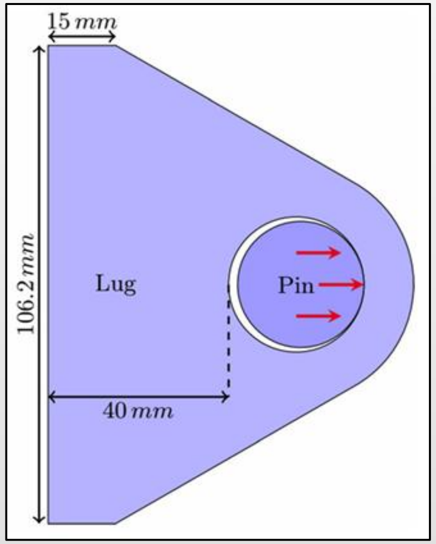
- **Contact between a cylinder and a rectangular block**
- **500 samples**
- **Accuracy $= 98.60\%$**



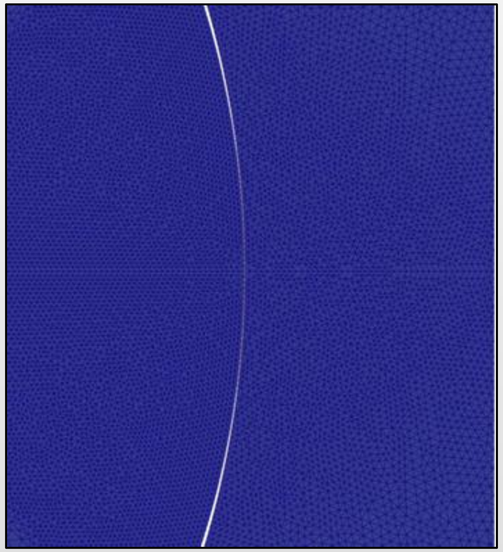
- **Contact between a cylinder and a hollow one**
- **500 samples**
- **Accuracy $= 95.40\%$**

Application: Pin-lug connection

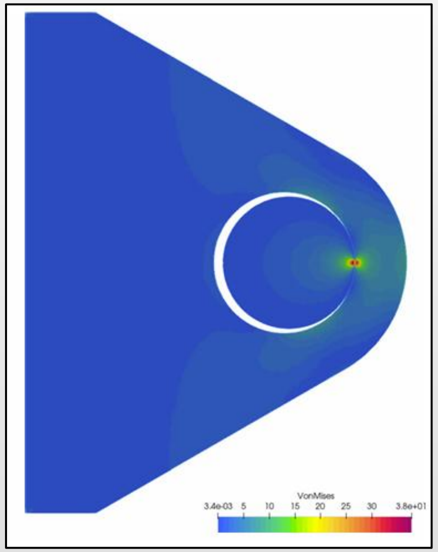
- Pin-lug connection composed of a lug (aluminum) and a pin (steel), typically used for structural lifting



Geometry

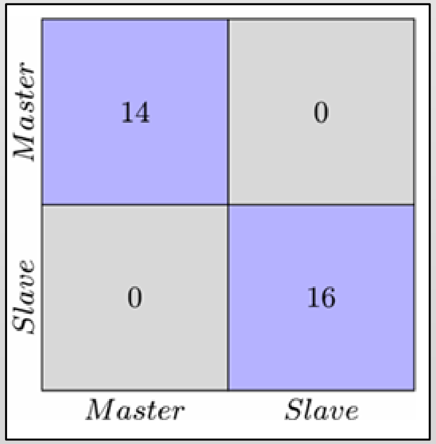


Meshes



Von Mises stresses

- Pin radius = 14mm and horizontal force of 50kN=m on the pin
- Hertzian contact (30 simulations)
- The ratio of the mesh density at the contact surface, r_n, varies randomly
- Recommendations fail to provide accurate guidance, in contrast to our ANN



Confusion matrix by the third ANN

Bibliography

- ❑ Johnson KL (1987) Contact Mechanics. Cambridge University Press
- ❑ Houssein H (2022) Finite element modeling of mechanical contact problems for industrial applications. PhD thesis, Sorbonne Université
- ❑ Hecht, F. (2012). New development in FreeFem++. *Journal of numerical mathematics*, *20*(3-4), 1-14.
- ❑ Houssein, H. An artificial neural network for predicting the roles of contacting bodies in computational contact mechanics. *Neural Comput & Applic* **37**, 27003–27025 (2025). https://doi.org/10.1007/s00521-025-11654-z
- ❑ Kingma DP (2014) Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980
- ❑ Chollet F, et al. (2015) Keras. https://keras.io
- ❑ Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mane´ D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Vie´gas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) TensorFlow: Large Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org. https://www.tensorflow.org/

THANK YOU FOR YOUR ATTENTION