Machine Learning (ML) in Physics

Dr. Tuhin Malik
Researcher, University of Coimbra,
Dept. of Physics, Coimbra, PT
Email: tuhin.malik@uc.pt

Date: 24 Nov 2021, 14:00 hr.

Outline

1 2 9 0 UNIVERSIDADE D COIMBRA

History

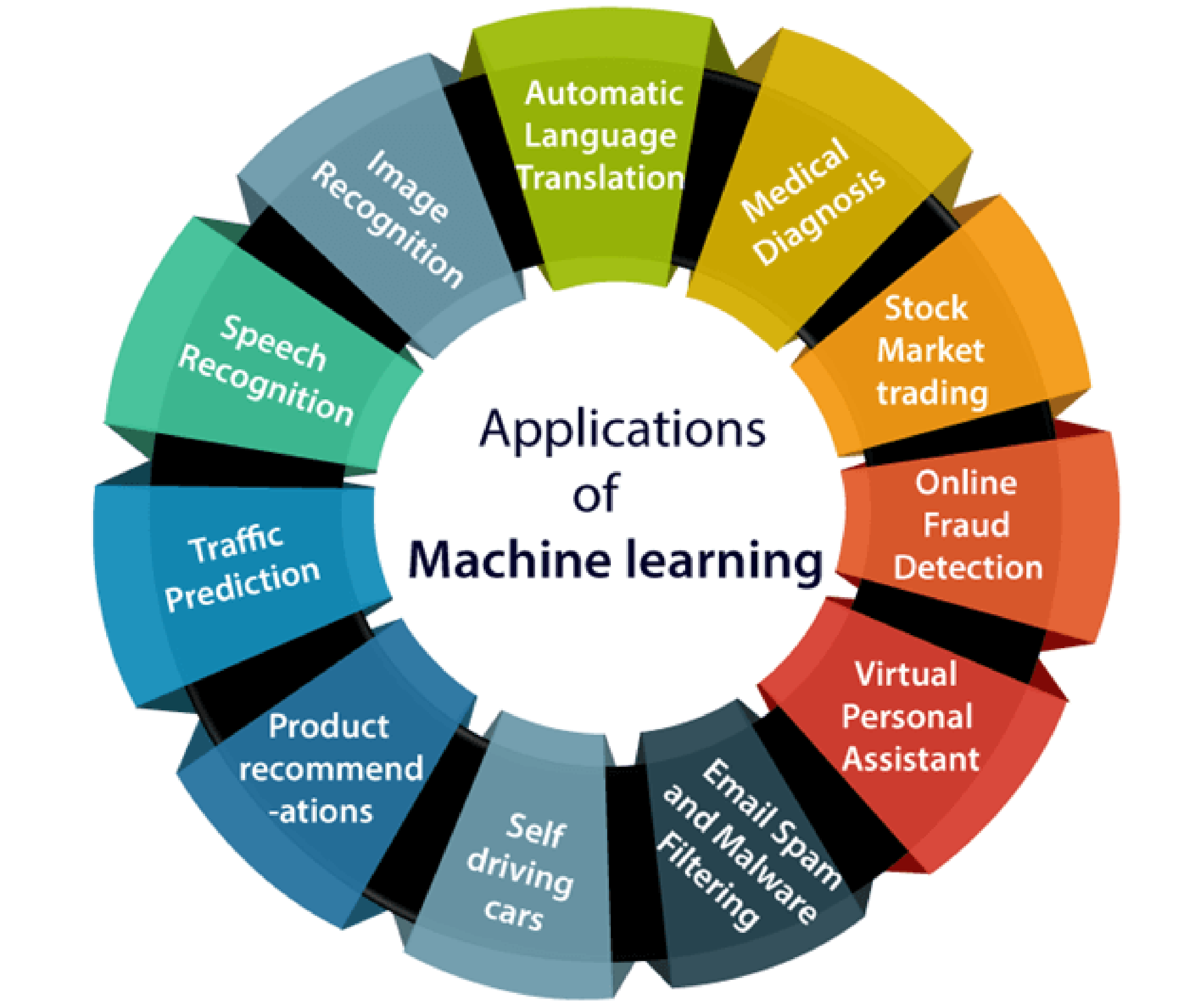
"A logical calculus of the ideas immanent in nervous activity"- McCulloch, W.S., Pitts, W. A, Bulletin of Mathematical Biophysics 5, 115–133 (1943), https://doi.org/10.1007/BF02478259

- Non-linear elements with weighted inputs have been suggested as simple models of biological neurons.

The Dartmouth Summer Research Project on Artificial Intelligence was a 1956 summer workshop widely considered to be the founding event of artificial intelligence as a field.

- Venue: Dartmouth College, Hanover, New Hampshire

1 2 9 0 UNIVERSIDADE D COIMBRA

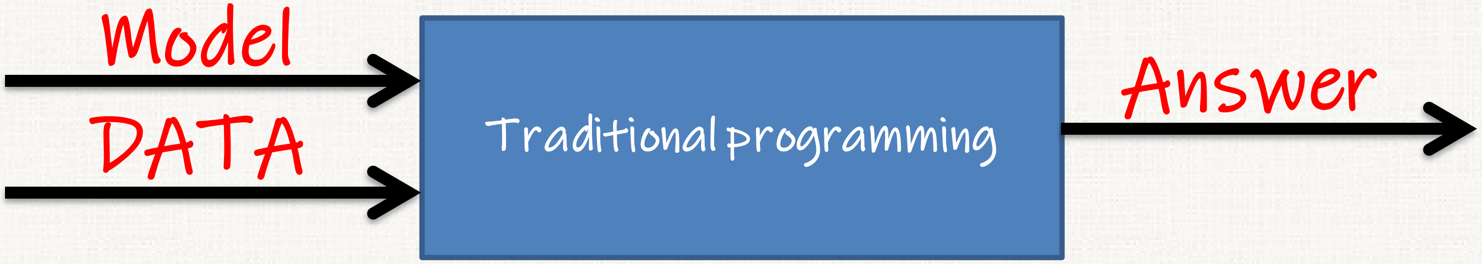


PC: www.javatpoint.com/applications-of-machine-learning

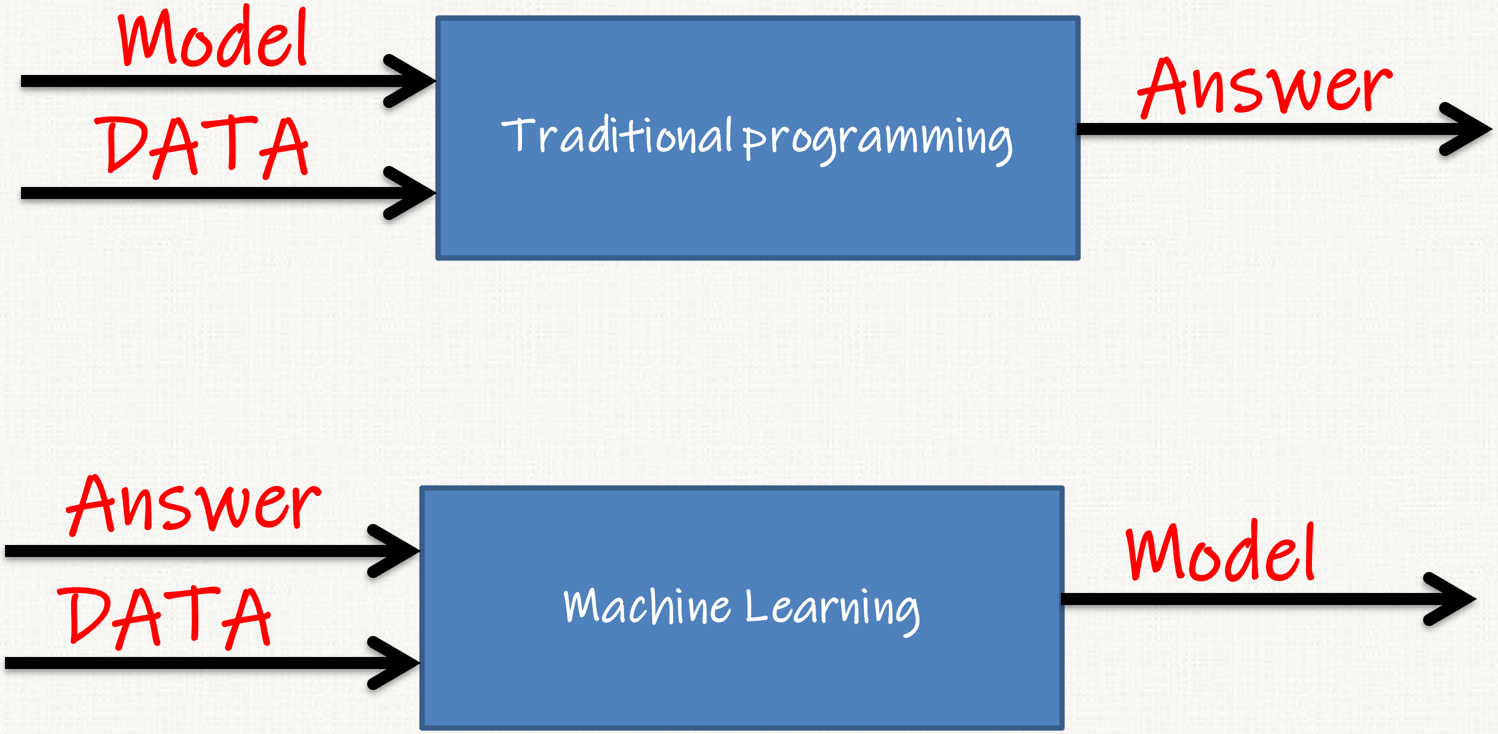
1 2 9 0 UNIVERSIDADE D COIMBRA

Use of Machine Learning in Physics:

- Machine learning for molecular and materials science (Nature 559, 547–555 (2018)) By Keith T. Butler, Daniel W. Davies, Hugh Cartwright, Olexandr Isayev & Aron Walsh
- Deep Learning and its Application to LHC Physics. (arXiv:1806.11484) By Dan Guest, Kyle Cranmer, Daniel Whiteson
- Deep Learning at Scale for Gravitational Wave Parameter Estimation of Binary Black Hole Mergers. (arXiv:1903.01998) By Hongyu Shen, E. A. Huerta, Zhizhen Zhao
- Real-Time Detection of Gravitational Waves from Binary Neutron Stars using Artificial Neural Networks. (arXiv:1908.03151) By Plamen G. Krastev (Harvard University)
- And many more

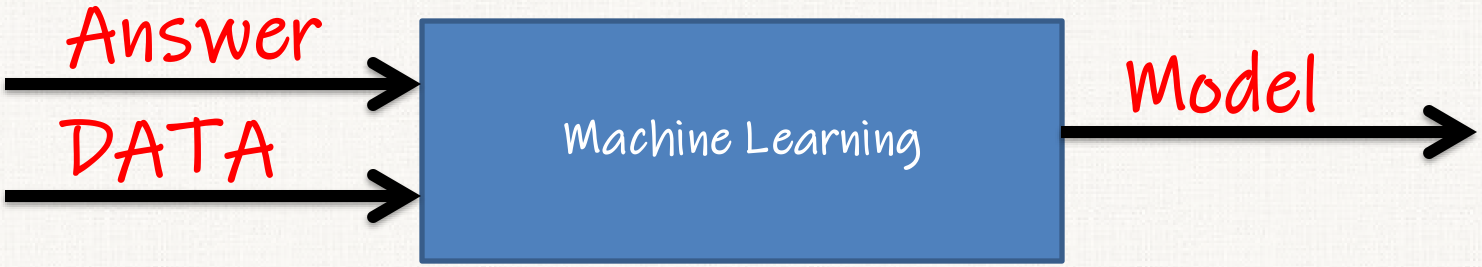


1 2 9 0 UNIVERSIDADE D COIMBRA

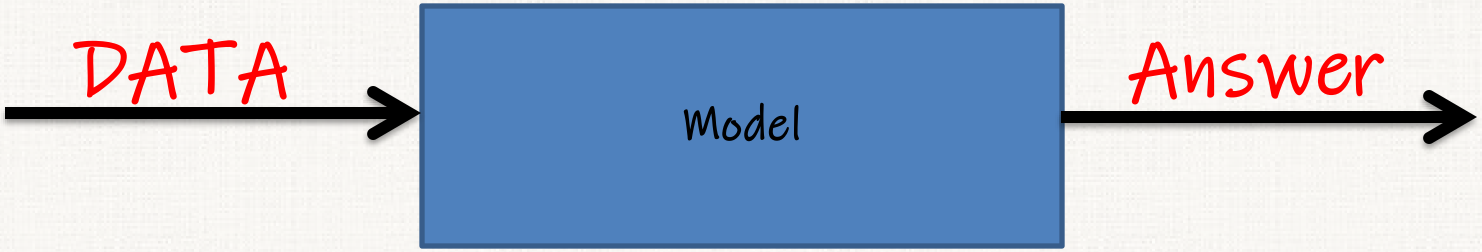


1 2 9 0 UNIVERSIDADE D COIMBRA
1 2 9 0 UNIVERSIDADE D COIMBRA

Training Phase



Inference Phase



1 2 9 0 UNIVERSIDADE D COIMBRA

y	2	6	10	14	?	?

1 2 9 0 UNIVERSIDADE D COIMBRA

y	2	6	10	14	?	?

y	2	6	10	14	18	22

1 2 9 0 UNIVERSIDADE D COIMBRA

x	0	2	4	6	7	9
y	2	6	10	14	?	?

1 2 9 0 UNIVERSIDADE D COIMBRA

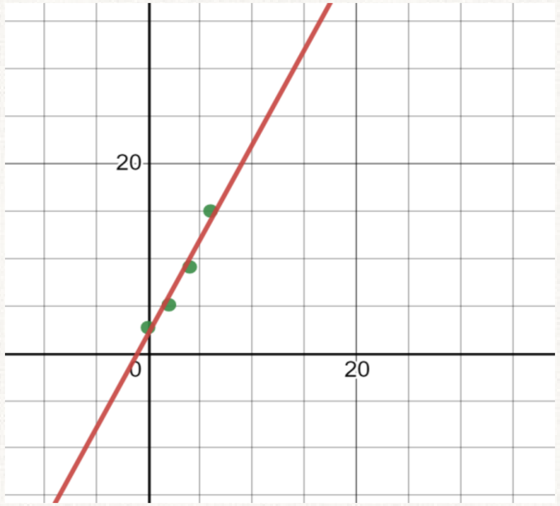
x	0	2	4	6	7	9
y	2	6	10	14	?	?

x	0	2	4	6	7	9
y	2	6	10	14	16	20

1 2 9 0 UNIVERSIDADE D COIMBRA

1 2 9 0 UNIVERSIDADE D COIMBRA

x	0	2	4	6	7	9
y	2.2	5.8	9.1	15	?	?



u=1.77
a=2.08

Model: $y = f(x)$

$$v = u + a\,t$$

x	0	2	4	6	7	9
y	2.2	5.8	9.1	15	16.33	20.49

1 2 9 0 UNIVERSIDADE D COIMBRA
1 2 9 0 UNIVERSIDADE D COIMBRA

Machine Learning

The aim of Machine Learning is to build a mathematical function which solve a human task.

Optimization

How to Learn ???

Learn from Data

"Universal functions"
"n" parameters

loss on target variable

Predict for New Data

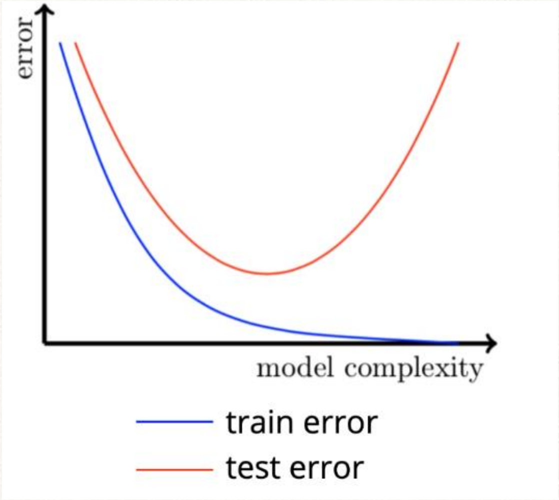
Data

Training set	Test set

Test set loss

Training subset	Validation Set

Tune Parameters

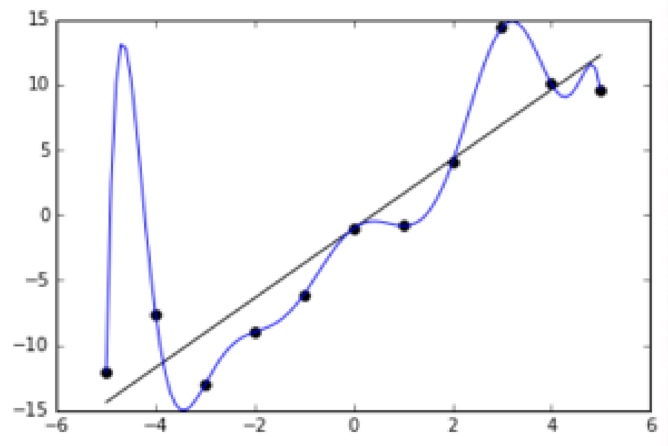


A simple loss function: $\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$

1 2 9 0 UNIVERSIDADE D COIMBRA

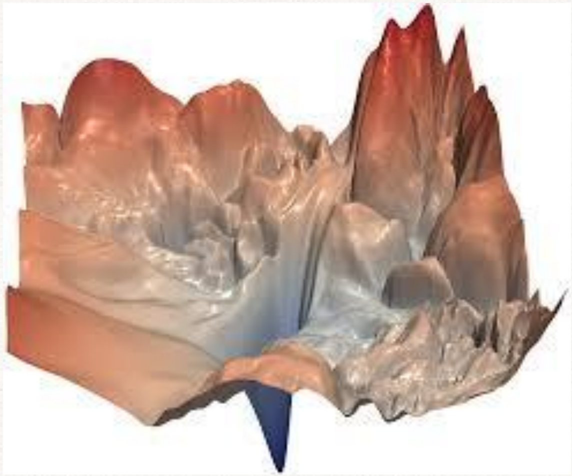
Challenges of ML

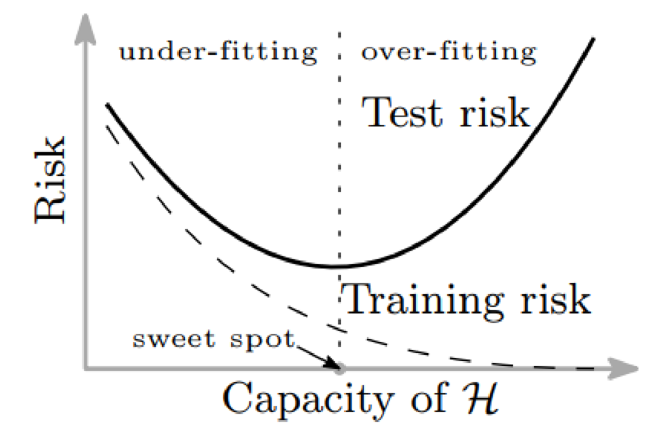
- Not enough training data.
- Poor Quality of data.
- Irrelevant features.
- Overfitting and Underfitting.



PC: https://en.wikipedia.org/wiki/Overfitting

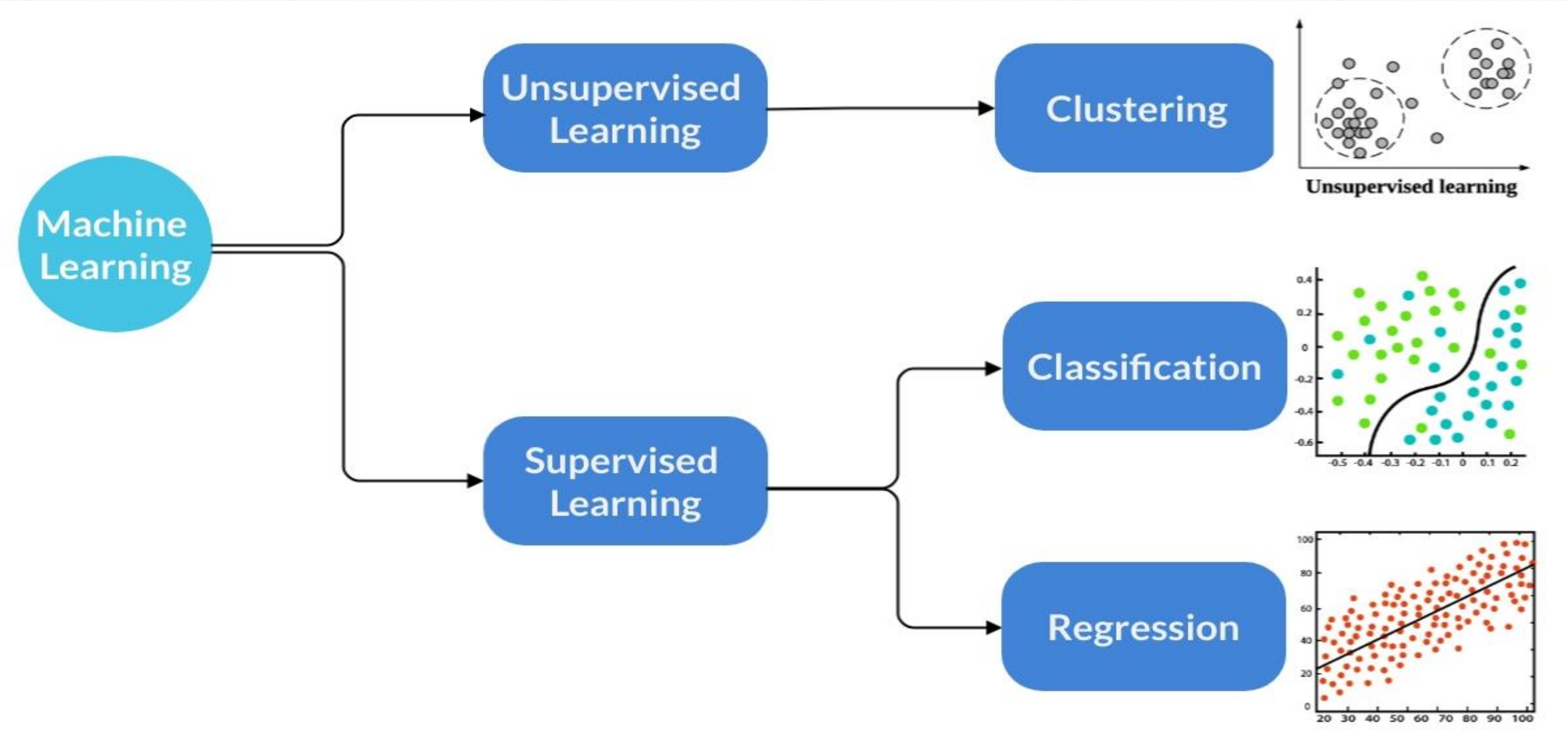
Global minima in the loss function





PC: https://blog.ml.cmu.edu/2020/08/31/4-overfitting/

1 2 9 0 UNIVERSIDADE D COIMBRA



- In supervised learning, the data that we have has two components, **features**, and **labels**.
- Labels are also known as the "target variable".
- In unsupervised learning, the data that we have is **unlabeled**.

1 2 9 0 UNIVERSIDADE D COIMBRA

Algorithm

- Decision Tree (DT)
- Random Forest (RF)
- Support Vector Machines (SVM)
- Stacked Generalization (SG)
- Neural Network (NN)

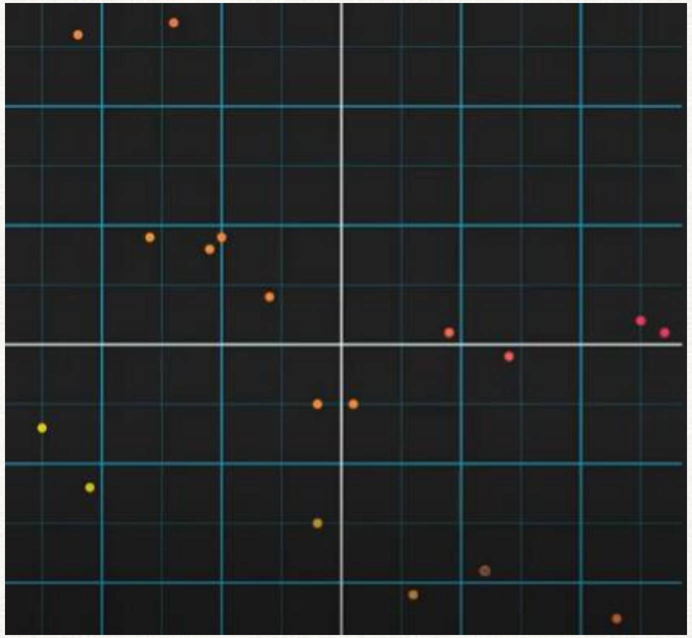
▪ There is no general procedure to choose the best ML algorithm suitable for a physics problem as it depends on the problem at hand as well as on the domain knowledge available for that problem.

▪ The choice of an ML algorithm depends heavily on the nature of the data set. Things like dimensionality of the data (number of features), number of data points, distribution of data, linearity of data, nature of features (continuous, discrete, categorical etc) influences which algorithm will best fit the data.

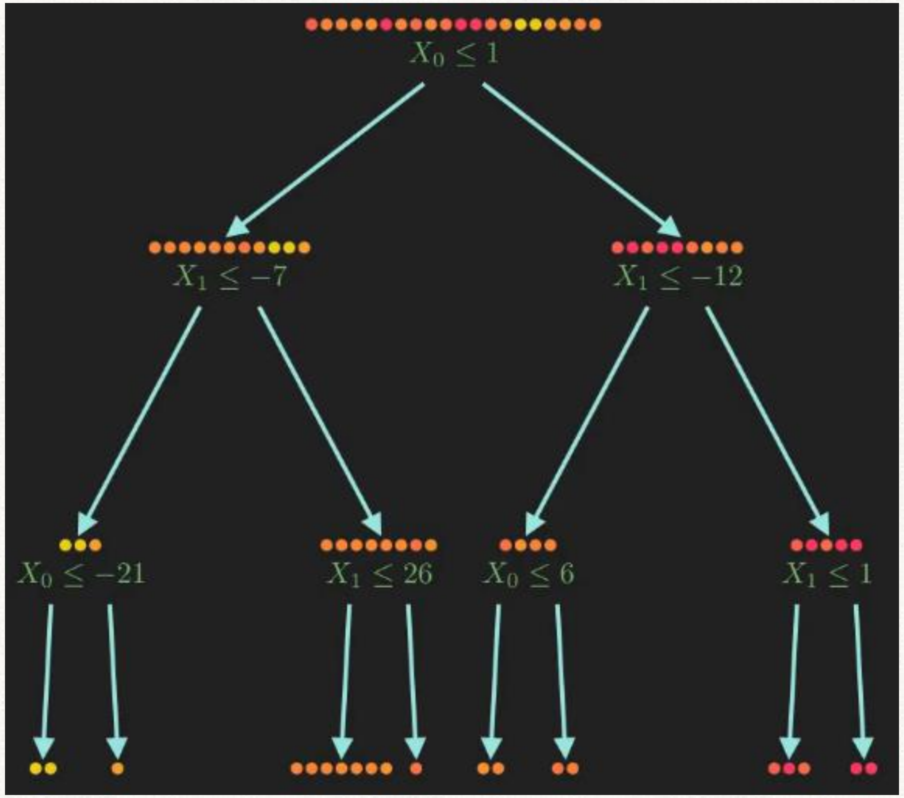
1 2 9 0 UNIVERSIDADE D COIMBRA

Decision Tree (DT)

Decision tree works by dividing the data set into smaller parts, which are similar in nature based on metrics like information entropy, variance, and impurity.



$$Var = \frac{1}{n}\sum(y_i - \bar{y})^2$$



$$Var\ Red = Var(parent) - \sum w_i Var(Child_i)$$

1 2 9 0 UNIVERSIDADE D COIMBRA

Random Forest (RF)-

Random Forests are an ensemble of decision trees. They generally perform better than decision trees because they average out the errors made by individual decision trees in the ensemble [1].

Support Vector Machines (SVM)-

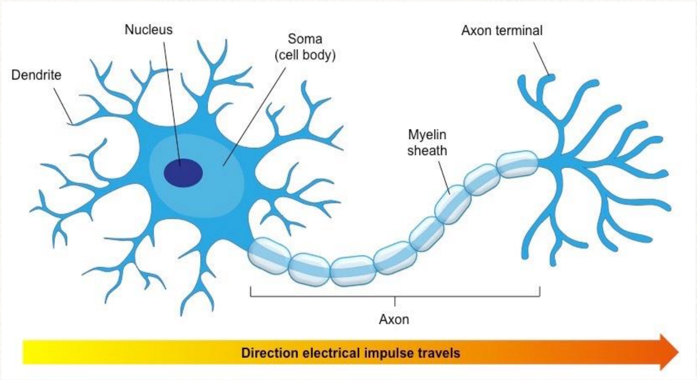
Support Vector Machines (SVM) make non-linear regression very easy because of the kernel trick. The SVM algorithm works by computing a similarity measure between two points in the feature space, we can define this similarity measure using a kernel which will effectively map all the points into a higher dimensional feature space, in which our data can be linear [2].

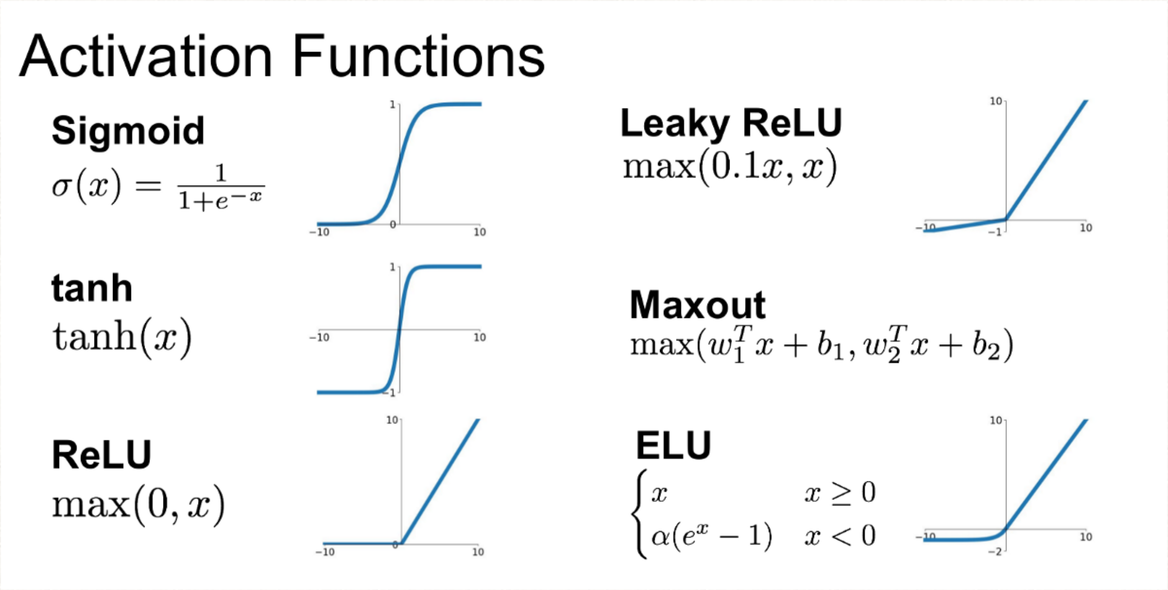
[1] Leo Breiman. Random forests. Mach. Learn., 45(1):5–32, October 2001. https://doi.org/10.1023/A:1010933404324.
[2] Harris Drucker, Christopher J. C. Burges, Linda Kaufman, Alex J. Smola, and Vladimir Vapnik. Support vector regression machines. In M. C. Mozer, M. I. Jordan, and T. Petsche, editors, Advances in Neural Information Processing Systems 9, pages 155–161. MIT Press, 1997. URL http://papers.nips.cc/paper/1238-support-vector-regression-machines.pdf.

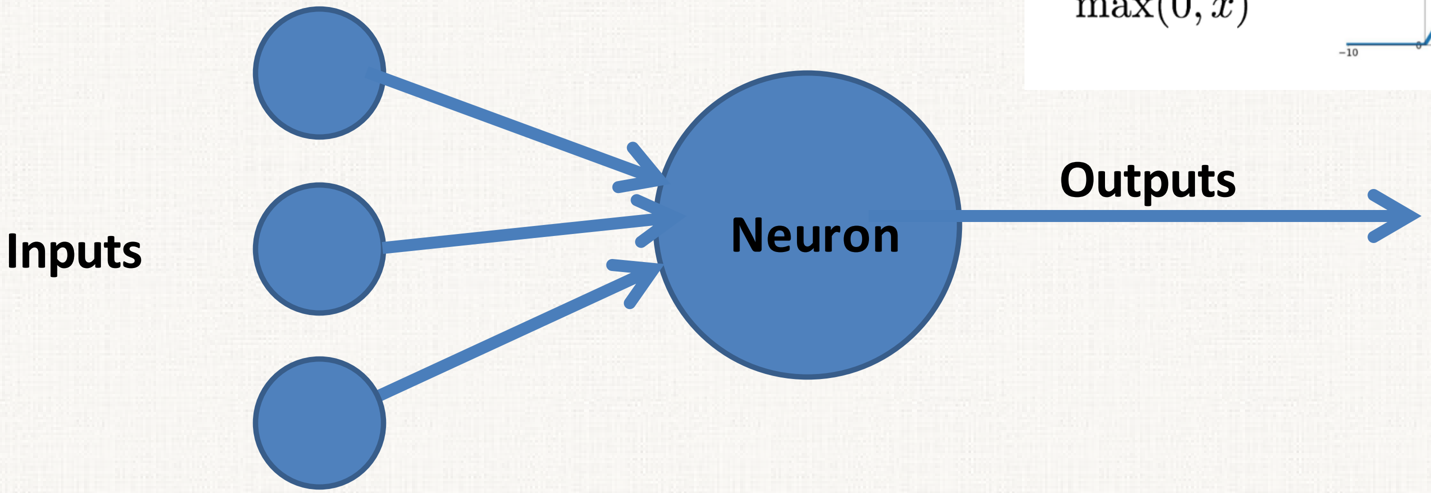
1 2 9 0 UNIVERSIDADE D COIMBRA

Neural Network

Neural networks are a class of machine learning algorithms that are loosely inspired by neurons in the human brain.



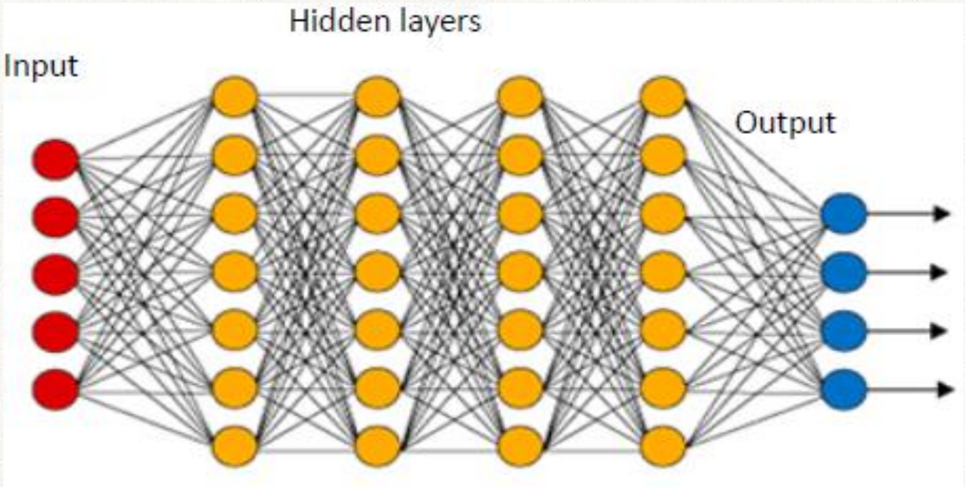




$$y = \phi(w.x + b)$$

1 2 9 0 UNIVERSIDADE D COIMBRA

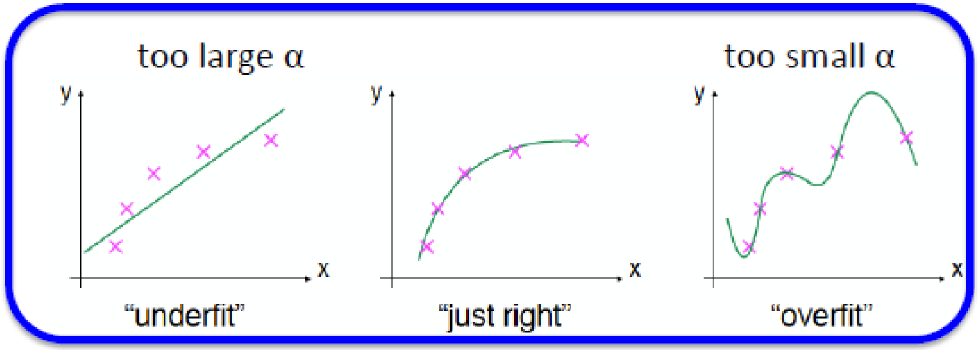
Deep Neural Network



Loss function: $L\left(\mathbf{W}\right)=\frac{1}{N_{\text{train}}}\sum_{t=1}^{N_{\text{train}}}\left(F\left(\mathbf{x}_i\right)-y_t\right)^2+\alpha\left\|\mathbf{W}\right\|_2^2$

mean squared error

regularization (penalizes large weights)



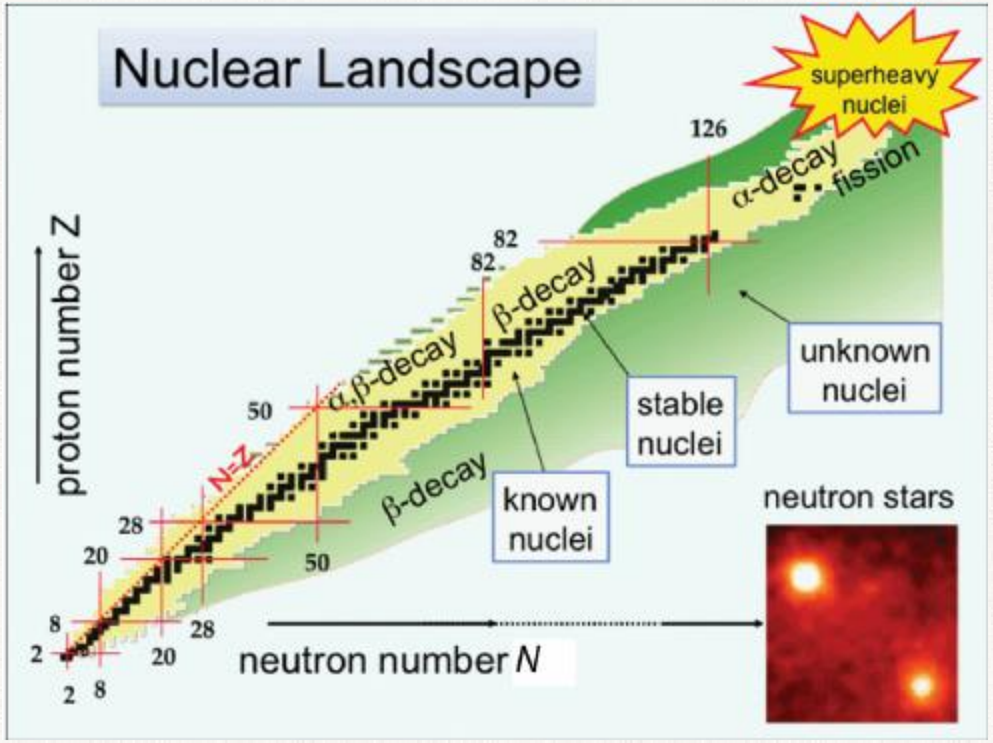
1 2 9 0 UNIVERSIDADE Ð COIMBRA

Physics Problem

- Nuclear Binding Energy (BE).
- Neutron Star matter (NS).

1 2 9 0 UNIVERSIDADE D COIMBRA

Binding Energy (BE) of the atomic nuclei



- ❑ The experimental measurements of nuclear BE have achieved a great success, in the atomic mass evaluation AME2016, 3435 nuclei have been measured in the laboratories around the world.
- ❑ Recently published AME2020 data: the 122 nuclei newly added.

- ✓ Many nuclei of astrophysical relevance still remain beyond the experimental reach.
- ✓ Thus, theoretical modeling of nuclear theory that extrapolate BE into unknown regions of the nuclear chart becomes very important.
- ✓ Unfortunately, theoretical modeling of nuclei to predict BE is challenging due to the uncertain theories of nuclear interaction and difficulties in the quantum many-body calculations.

We do not assume any nuclear physics models in our work. We predict nuclear BE based only on Machine Learning algorithms. Our data set is AME2016, where we only look at Z, N and BE data of all nuclei.

- ❖ We break our data set randomly into 60% training data 20% validation and 20% testing set. In the first part we chose the base algorithm of ML which gives the lowest RMSE.
- ❖ In the second part we train the error obtained from the base algorithm to further reduce the RMSE on the test set.
- ❖ The RMSE for each ML model is stabilized by doing multiple runs with randomly chosen training, validating, and testing sets.

1 2 9 0 UNIVERSIDADE D COIMBRA

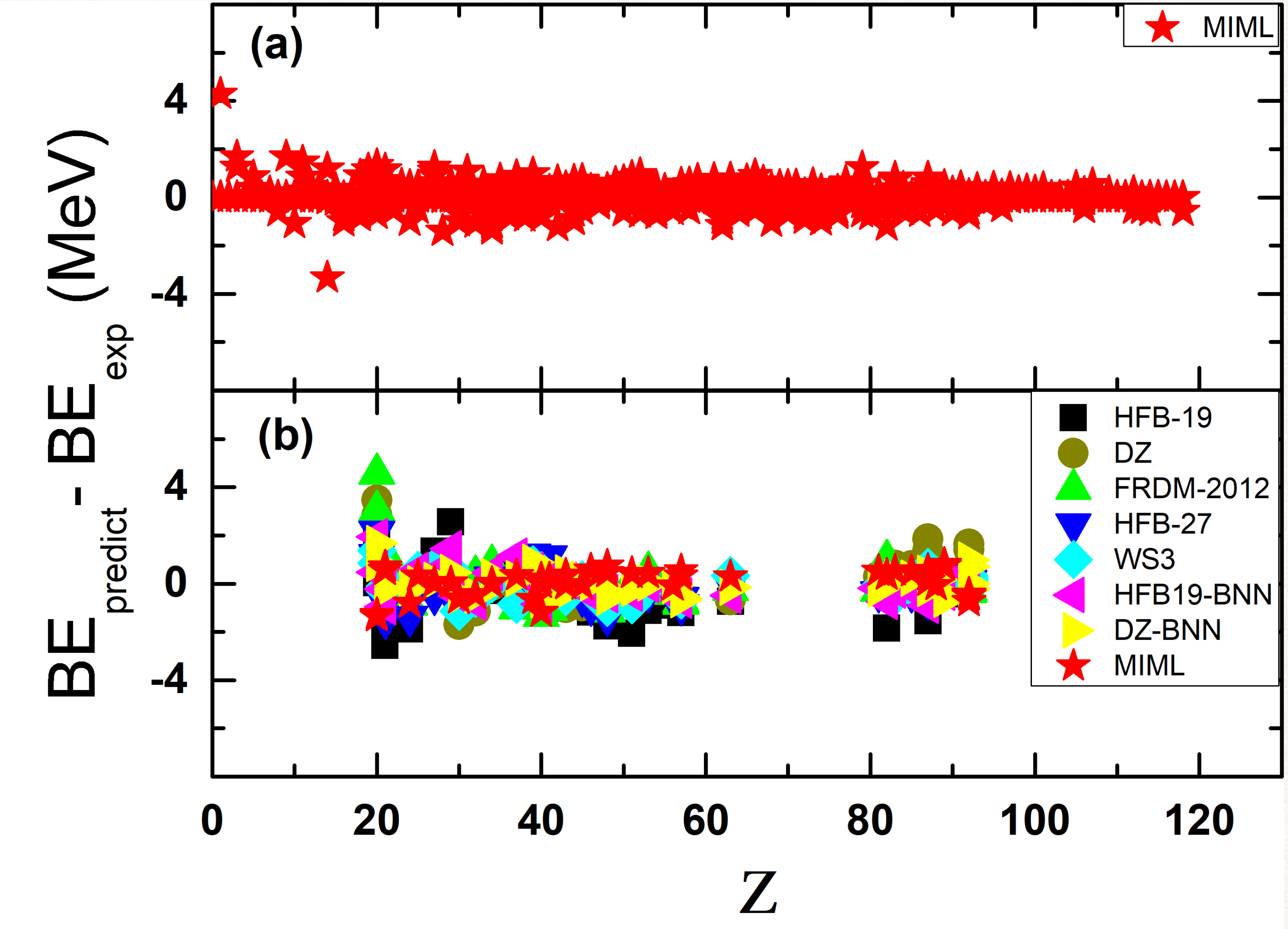
Table 1. The $\sigma_{\rm rms}$ error of base ML models ($\sigma^i_{\rm rms}$) as well as error trained model ($\sigma^f_{\rm rms}$) is outlined. n Random forest models on top of the base model are considered for error trained model. All the errors are for the test set only

Base Model	$\sigma^i_{\rm rms}$ (MeV)	n	$\sigma^f_{\rm rms}$ (MeV)
LR	45.78	3	1.65
DT	8.22	6	6.32
RF	2.18	4	1.56
PR	2.58	4	1.18
SVM	1.81	10	0.58

Table 3. $\sigma_{\rm rms}$ of the predictions of various models for the 46 nuclei in the ^{40}Ca - ^{240}U region that appear in the latest AME2016 [8] compilation but not in AME2012 [31].

Model	HFB-19	DZ	FRDM-2012	HFB-27	WS3	HFB19-BNN	DZ-BNN	MIML
$\sigma_{\rm rms}$	1.093	1.018	0.997	0.723	0.513	0.587	0.479	0.501

https://arxiv.org/abs/2004.14196



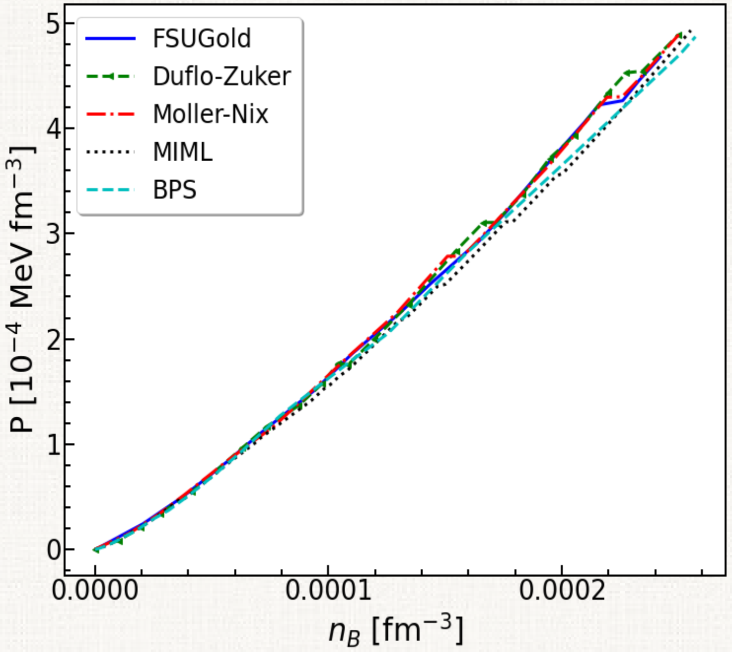
1 2 9 0 UNIVERSIDADE D COIMBRA

Test on New Data !!

AME2020: The 122 nuclei newly added. The RMSE obtained in MIML for the entire AME2020 data is 0.2224 MeV, and for the newly added 122 is 0.7772 MeV.

The composition of the NS outer crust can be determined by minimizing the Gibbs free energy or chemical potential at a certain pressure with respect to the atomic mass number A.

- Several nuclei appear with an odd number of protons or neutrons.
- This seems a weakness of the MIML model which is not able to capture the pairing effect fully.



The comparison of the pressure as a function of baryon density for the NS outer crust matter for different model.

1 2 9 0 UNIVERSIDADE D COIMBRA

The Neutron Star

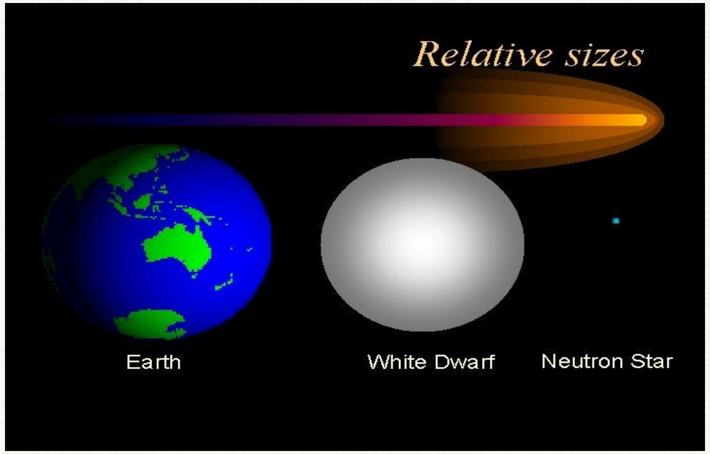
1 2 9 0 UNIVERSIDADE D COIMBRA
1 2 9 0 UNIVERSIDADE D COIMBRA

The NS properties

$M_\odot$ = 1.989 x 10^{30} kg $R_\odot$ = 696340 km

The observation of high mass pulsars

$$PSR\,J1614-2230\,(M = 1.928\ \pm 0.04\ M_\odot)[1]$$
$$PSR\,J0348-0432\,(M = 2.01\ \pm 0.04\ M_\odot)[2]$$
$$PSR\,J0740+6620\,(M = 2.08\ \pm 0.07\ M_\odot)[3]$$
$$PSR\,J1810+1714(M = 2.13\ \pm 0.04 M_\odot)[4]$$



The NICER radius measurement

$$PSR\,J0740+6620\ (R = 13.39^{+1.30}_{-0.98}\ km)[5]$$

The empirical estimates $(R_{1.4} = 11.9^{+1.22}_{-1.22}\ km)$[6]

- Central no. density 4 – 8 ρ_0
- Asymmetry $I = \dfrac{\rho_n - \rho_p}{\rho_n + \rho_p} \approx 0.7$

Ref,

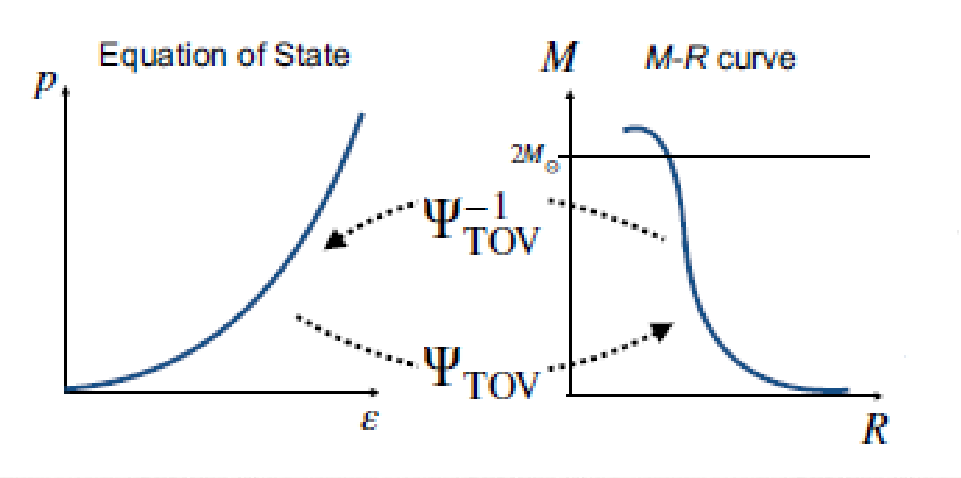
[1] Astrophys. J. 832, 167 (2016) , [2] *Science* 340 (2013) 6131
[3] Astrophys. J. Lett. 915, L12 (2021), [4] Astrophys. J. Lett. 908, L46 (2021)
[5] Astrophys. J. Lett. 918, L27 (2021), Astrophys. J. Lett. 918, L28 (2021), [6] Astrophys. J. 771, 51 (2013)

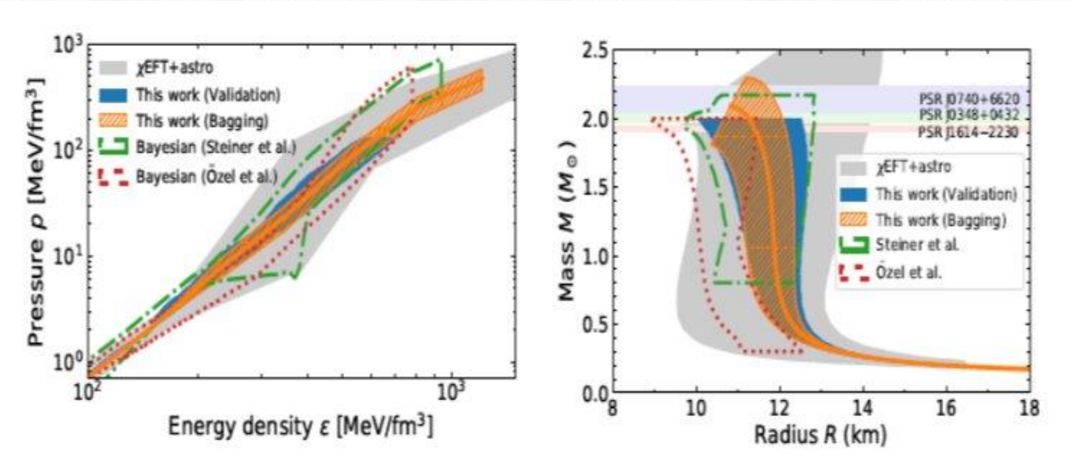
Structure equations (Spherical case)

Tolman-Oppenheimer-Volkoff (TOV) equations

$$\frac{dP(r)}{dr} = -\frac{(\epsilon(r)+P(r))\,(m(r)+4\pi r^3 P(r))}{r(r-2m(r))}$$

$$\frac{dm(r)}{dr} = 4\pi r^2 \epsilon(r)$$





JHEP 03 (2021) 273

1 2 9 0 UNIVERSIDADE D COIMBRA

The equation of state (EoS)

$$p(\rho) = \rho^2 \frac{d}{d\rho}(e(\rho))$$

The energy per particle at a given density ρ and asymmetry δ can be decomposed, to a good approximation

$$e(\rho, \delta) = e(\rho, 0) + S(\rho)\, \delta^2 \qquad \rho = \rho_n + \rho_p \qquad \delta = \frac{\rho_n - \rho_p}{\rho}$$

Symmetry Energy

SNM

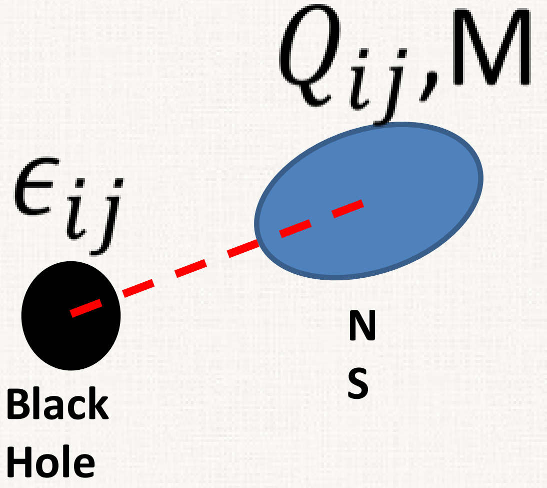
$$S(\rho) = J_0 + L_0 \left(\frac{\rho - \rho_0}{3\,\rho_0}\right) + K_{sym,0} \left(\frac{\rho - \rho_0}{3\,\rho_0}\right)^2 + \mathrm{O}(3)$$

$$e(\rho, 0) = e(\rho_0) + \frac{K_0}{2} \left(\frac{\rho - \rho_0}{3\rho_0}\right)^2 + \frac{Q_0}{6} \left(\frac{\rho - \rho_0}{3\rho_0}\right)^3 + O(4)$$

$$M_0 = Q_0 + 12 K_0$$

1 2 90 UNIVERSIDADE D COIMBRA
1 2 90 UNIVERSIDADE D COIMBRA

The tidal deformability



Quadrupole Moment

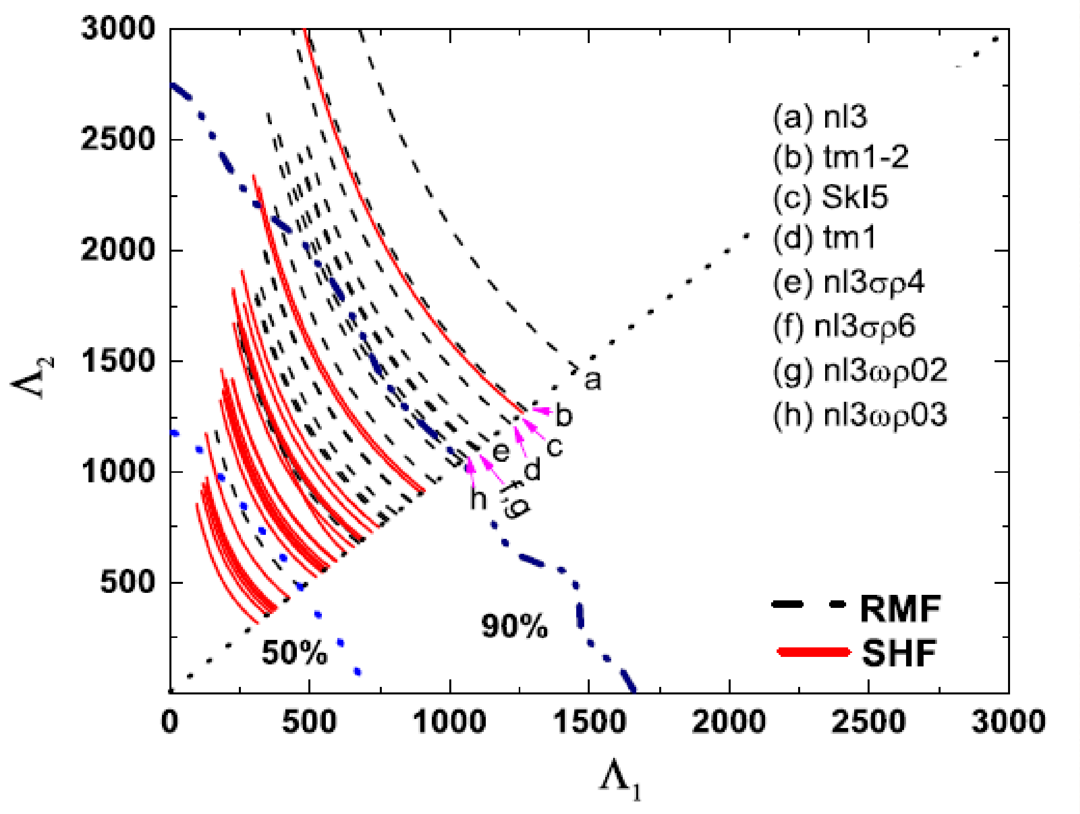
Tidal field

$$Q_{ij} = \lambda\, \epsilon_{ij}$$

$$\lambda = \Lambda\, M^5$$

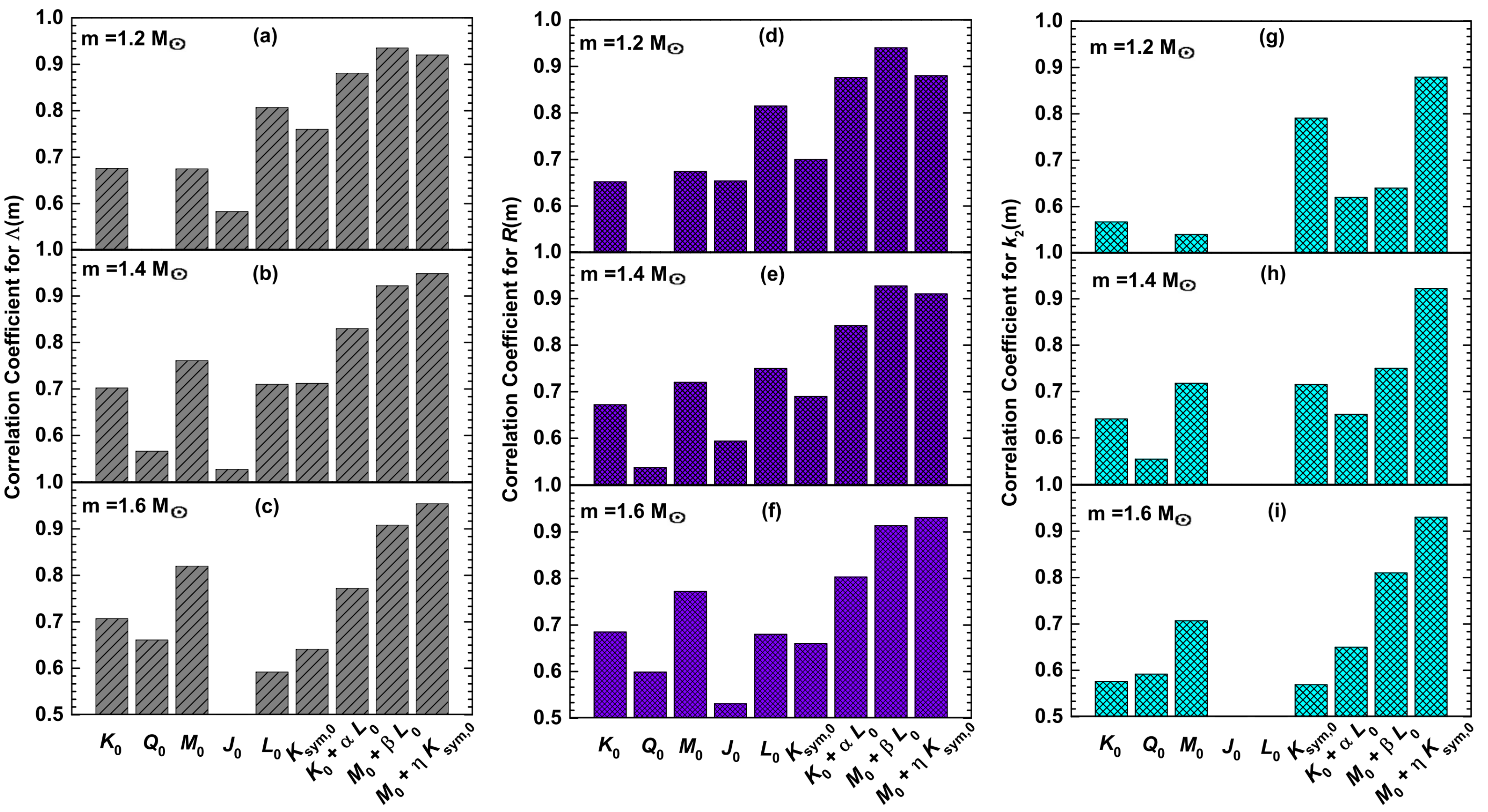
$$\Lambda = \frac{2}{3}\, k_2\, R^5$$

GW170817



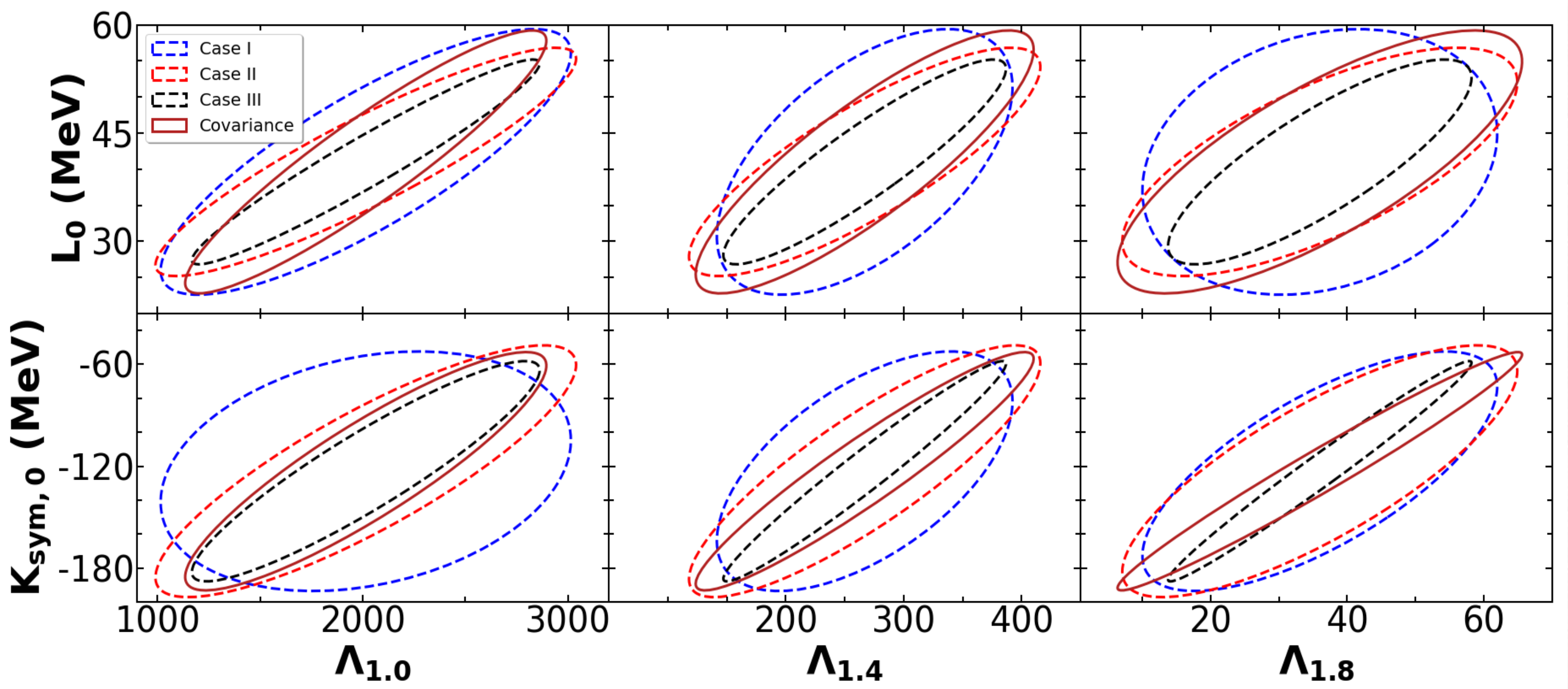
Malik et al, *PRC* 98 (2018) 3, 035804

1 2 9 0 UNIVERSIDADE D COIMBRA

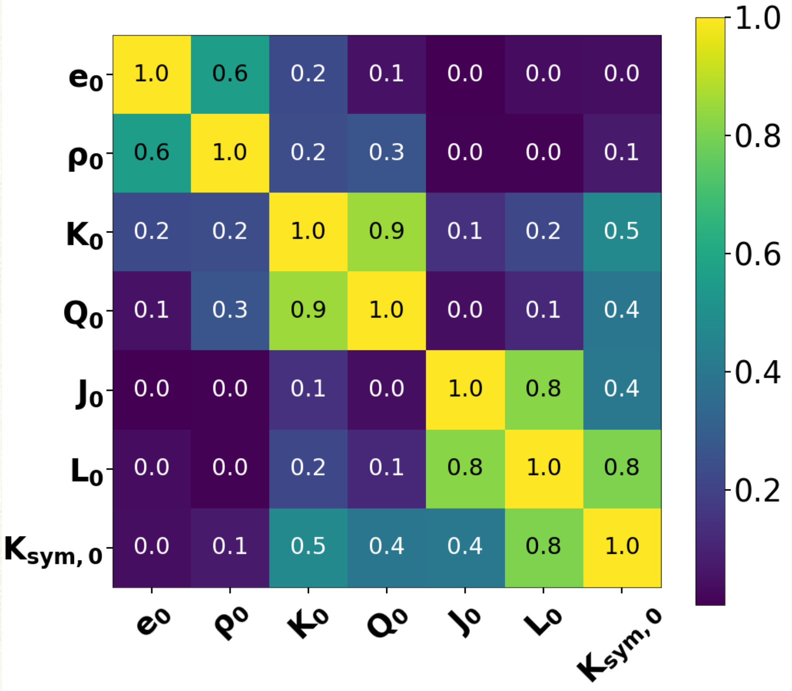


UNIVERSIDADE D COIMBRA
UNIVERSIDADE D COIMBRA

Unveiling the correlations of tidal deformability with the nuclear symmetry energy parameters



The 1σ confidence ellipses in the planes of $\Lambda_M - L_0$ (top) and $\Lambda_M - K_{sym,0}$ (bottom) with M=1.0, 1.4, and $1.8\,M_\odot$ obtained for Cases I-III and Sk Λ267. The central values of all the quantities for all cases are matched to those for Sk Λ267 for the appropriate comparison. The actual central values for these cases are L_0=60, $K_{sym,0}$=−100, $\Lambda_{1.0}$=3200, $\Lambda_{1.4}$=430, and $\Lambda_{1.8}$=60.



Correlations among various NMPs obtained using 237 selected Skyrme models from Ref. [53]. The correlations among the off-diagonal pairs $K_0 - Q_0$, $J_0 - L_0$, and $L_0 - K_{sym,0}$ are noticeable.

Malik et al, PRC C 102, 052801(R) (2020)

1 2 9 0 UNIVERSIDADE D COIMBRA

Neural network Model

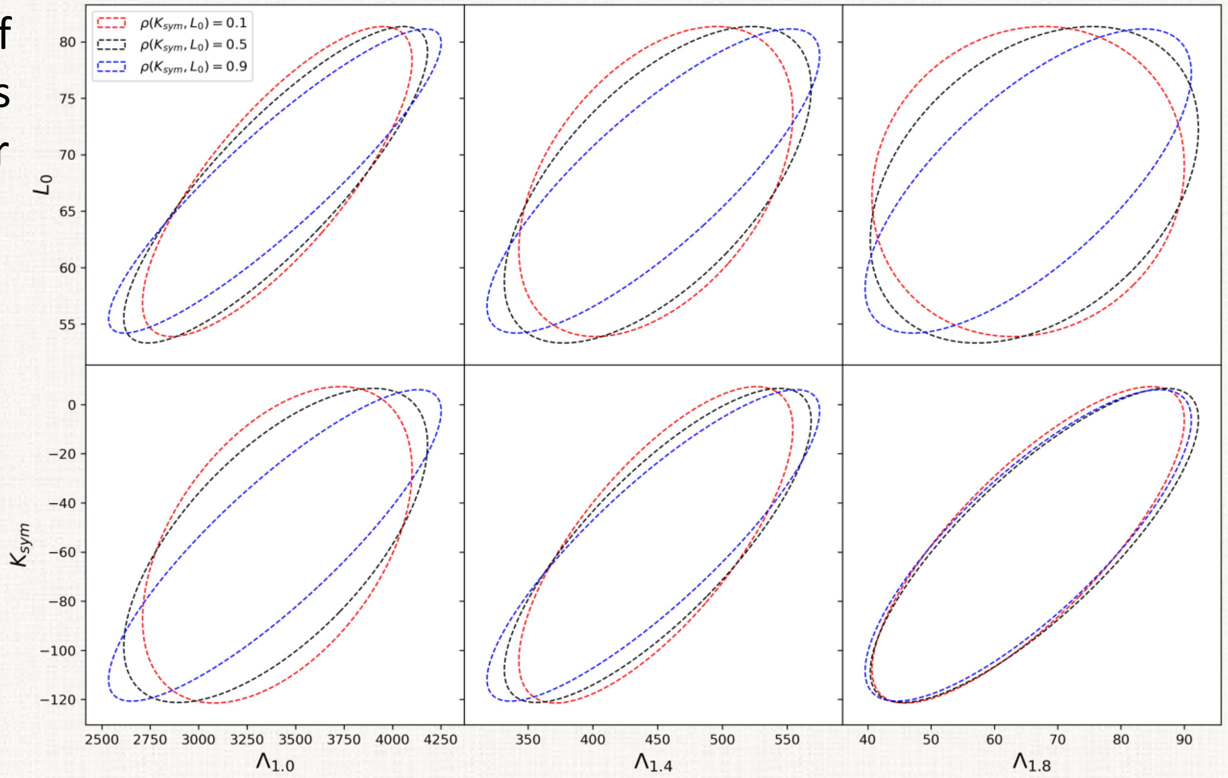
We adopted a neural network architecture consisting of two hidden dense layers consisting of fifteen neurons each and a final output layer consisting of 6 neurons for the target properties.

features

$$e_0, \rho_0, K_0, Q_0, J_0, L_0, K_{\mathrm{sym},0}$$

target variable

NS Property	RMSE
NS Mass	0.0243
$R_{\max}$	0.0879
$R_{1.4}$	0.1940
$\Lambda_{1.0}$	123.8
$\Lambda_{1.4}$	19.23
$\Lambda_{1.8}$	5.344




```
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 15)                120

dense_1 (Dense)              (None, 15)                240

dense_2 (Dense)              (None, 6)                 96
=================================================================
Total params: 456
Trainable params: 456
```


1 2 9 0 UNIVERSIDADE D COIMBRA

Conclusions

- The role of theoretical models go way beyond producing numbers. A theoretical model also indicates the actual physical mechanisms behind the properties being predicted. Since each term in the model is physically motivated, a theoretical model which comes close to experimental predictions also identifies what are the actual physical processes which are important in that energy scale. To have a theoretical understanding of any system, a physics based model is necessary. ML algorithms cannot replace physics modeling in that respect.

- However an interesting area of future work might be in combining the theoretical model and the machine learning methods to arrive at a better physical models. Our theoretical knowledge may help determine which features are physically relevant in a given data set while ML algorithms will help us find patterns and make predictions.

- Randomized Testing (**Testing....Testing....Testing**)

1 2 9 0 UNIVERSIDADE D COIMBRA

GOA, India

1 2 90 UNIVERSIDADE D COIMBRA