Exploring multi-fidelity bayesian neural network for nuclear data evaluation

Aman Sharma,* A. Gandhi, and Ajay Kumar
Department of physics, Banaras Hindu University, Varanasi 221005, India

Introduction

Nuclear reaction data along with uncertainty quantification is of primary interest for nuclear science and technology. The process of measuring the nuclear reaction data over a broad spectrum of energy is expensive and difficult task if not impossible in some cases [1–4]. Therefor theoretical models play a crucial role in the field of nuclear data. The theoretical predictions and available experimental data of good quality is used for the evaluation of various nuclear reaction observables. These final evaluations are then used by the end users.

There have been a number of attempts in past few years for utilising the potential of machine learning in nuclear science [5, 6]. In this study we have explored the possibility of using bayesian neural networks for the nuclear data evaluations. This study is motivated by the fact that low fidelity data from theoretical models is ample, while the high fidelity data from experiments is scarce. We have used a multi-fidelity bayesian neural network to benefit from both of these sources of the data [7]. One of the benefits of using bayesian neural network is that they are less prone to over fitting then normal neural networks. Because in general neural networks the model learns fixed values of the wights and the biases, but in the bayesian neural network the model assigns distributions to the weights and biases and these distributions are optimised using the bayesian inference. Hence bayesian neural networks are able to provide the distribution of the final outcome, from which statistical moments of the final outcomes can be calculated. In this study we have evaluated the differential cross sections of the elastic scattered neutrons from ^{238}U.

*Electronic address: aman.marley1314@gmail.com

Methodology

The present model consists of a feed forward neural network, which learns the low fidelity data (LFNN) and acts as a surrogate model to the low fidelity theoretical simulations. Then the output of LFNN with the high fidelity data inputs are feed in to a bayesian neural network (HFBNN) which is optimised with respect to the high fidelity data. A schematic representation of the methodology is presented in the Fig. 1.

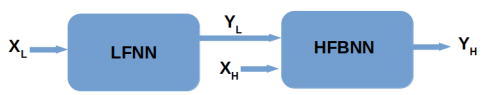

FIG. 1: A schematic representation of the present model.

Here X_L and Y_L represents respectively the low fidelity inputs and outputs which may be from theoretical models. X_H and Y_H represents the high fidelity inputs and outputs, also $\hat{Y}_L$ which are the outputs of LFNN corresponding to the X_H are also taken as the inputs in HFBNN. In this study we have use the elastic scattering differential cross section data of the neutrons from the ^{238}U at energies 14.2 MeV as a test case. Neutron scattering lab angles are taken as the inputs and the elastic differential cross sections as the output values. The elastic differential cross sections corresponding to 1800 angles (0^0 to 180^0) were calculated using TALYS nuclear reaction code and then were used to train LFNN. We have used a 1-50-50-50-1 neural network as LFNN,

with 'elu' activation function. For regularisation we have used L2 regularisation with regularisation rate of 0.0001. The weights and biases were optimised by using the Root Mean Squared Propagation algorithm with learning rate 0.001. Training and validation sets were taken to be 80% and 20% respectively.

We have used 48 experimental values of the differential cross sections for the present reaction from the EXFOR data library and considered it as the high fidelity data. LFNN predictions corresponding to the experimental angles were calculated and then used as the input to the HFBNN along with the lab angles. We have used only one hidden layer in the bayesian neural network with 10 nodes. We have used multivariate normal distributions as our prior and posterior distributions for the hyper-parameters of the neural network. Both of the neural networks were optimised using the mean squared error as the loss function.

Results

Experimental data along with the Talys predictions and predictions from present study is presented in Fig.2. Here the black curve presents the TALYS predictions for the elastic differential cross sections at neutron energy 14.2 MeV. The blue dots represents the experimental data and the red curve presents the mean of the elastic differential cross sections predictions from the present model. The shaded region around the red curve depicts the 95% confidence interval of the present model predictions. The mean and confidence intervals were calculated by using 1000 outputs of the model. It is clear from Fig.2 that the present model has successfully learned the correlations between the low and high fidelity data. In this study, we have taken in to account the epistemic uncertainties that is the uncertainties in the energy region where no experimental data was available. We are also presently working on incorporating the aleatoric uncertainties in our model (the uncertainty due to the error bars in the experimental data). Also since we have directly used the theoretical model prediction as the inputs to our bayesian neural networks hence extrapolations of the data will be more reliable.

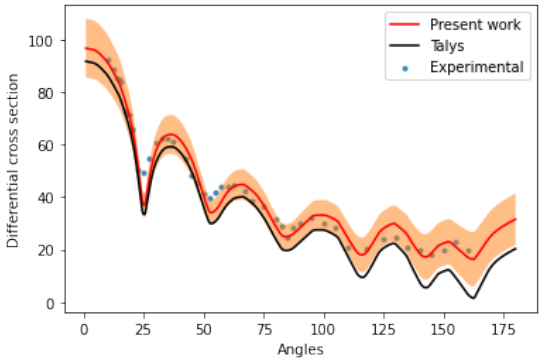

FIG. 2: Elastic differential cross sections for n+ ^{238}U at energies 14.2 MeV from EXFOR, TALYS and present study.

Acknowledgement

A. Kumar would like to acknowledge the financial support from SERB-DST, Government of India (Sanction No. CRG/2019/000360), IUAC-UGC, Government of India (Sanction No. IUAC/XIII.7/UFR- 71353) and Institutions of Eminence (IoE) BHU [Grant No. 6031] for this work.

References

[1] Aman Sharma *et al.*, Physics Letters B, 815 (2021) 136179
[2] A. Gandhi *et al.*, The European Physical Journal Plus, 136(8) (2021) 1 − 17
[3] A. Gandhi *et al.*, Chinese Physics C, 46(1) (2022) 014002
[4] Aman Sharma *et al.*, Physical Review C, 105 (2022) 014624
[5] Aman Sharma *et al.*, Physical Review C, 105 (2022) **L**031306
[6] G. E. Karniadakis *et al.*, Nature Rev Phys 3, 422440 (2021)
[7] X. Meng *et al.*, Journal of Computational Physics 438 (2021) 110361