Comparative Analysis of Nuclear Models with Machine Learning

Kunal Singh(UID-219114, Roll No.-9)
Tanishka Manjrekar (UID-219128, Roll No.-19)

Abstract: *The value of Binding Energy of nuclei is important for the study of various processes and phenomena. Nuclear models are designed to predict Binding Energy of nuclei for which experimental values cannot be obtained. This paper proposes a Machine Learning based Nuclear Model and compares it with the Liquid Drop Model to test the efficacy of Machine Learning algorithms in this field. We also propose a hybrid model by building our Machine Learning model on top of the Liquid Drop Model.*

Keywords: Machine Learning, Nuclear Models, SEMF, Nuclear Models, Binding Energy, Liquid Drop Model, Bethe–Weizsäcker, AME2020

Theoretical Background and Review of Relevant Literature:

After the discovery of nucleus in the Atom it has been a fundamental task to explain the structure of nucleus. One of the simplest models which explain the structure of nucleus is Liquid Drop Model. It treats the nucleus as a drop of incompressible fluid of very high density, held together by the nuclear force. The model has a similarity to the structure of a spherical liquid drop. The semi-empirical mass formula (SEMF) (sometimes called Bethe–Weizsäcker mass formula) describes the liquid drop model. According to it the binding energy of the nuclei is:

$$\boldsymbol{E_B = a_v A - \frac{a_s A^{\frac{2}{3}}}{3} - \frac{a_c Z(Z-1)}{A^{\frac{1}{3}}} - \frac{a_A (N-Z)^2}{A} + \delta(N,Z)}$$

$\delta(N,Z) = +\delta_0 \quad Z, N\ even$

$= 0\ either\ Z\ or\ N\ odd$

$= -\delta_0 \quad Z, N\ odd$

$a_v\ is\ the\ volume\ term,$

$a_s\ is\ the\ surface\ tem,$

$a_c\ is\ the\ coulomb\ term,$

$a_A\ is\ the\ asymmetric\ term\ and$

$\delta(N.Z)\ is\ the\ pairing\ term$

The Binding Energy is one of the fundamental properties of atomic nuclei and most of the other properties of atomic nuclei like mass, decay lifetimes and reaction rates are governed largely by the Binding Energy.

In recent years, the experimental measurements of nuclear BE have achieved a great success, in the atomic mass evaluation AME2020 , 3558 nuclei have been measured in the laboratories around the world.

Many nuclei of astrophysical relevance remain out from experimental reach. It becomes very important to extrapolate them based on nuclear model. Liquid drop only considers macroscopic and its predictions diverge from experimental value.

Machine learning is a very powerful tool used to predict patterns and it has been used in the field of Nuclear physics too. Following is an overview of the research done on application of machine learning in nuclear models

In the paper *'A new method to improve the generalization ability of neural networks: A case study of nuclear mass training'(Tianliang Zhao, Hongfei Zhang)'* the authors use neural networks to improve the accuracy of the mass models like liquid drop model and Weizsäcker-Skyrme (WS) model and saw improvement in the accuracy.

In *'Machine learning the nuclear mass'(Zepeng Gao, Yongjia Wang, Hongliang Lü, Qingfeng Li, Caiwan Shen, Ling Liu)'* the authors use Light Gradient Boosting Machine (LightGBM) , a machine learning algorithm to refine predictions of nuclear model.

'A Study of Nuclear Structure and Neutron Stars with a Bayesian Neural Network Approach'(Raditya Utama)' in their paper suggest a hybrid model where they used Bayesian neural network to optimize theoretical mass models.

In our project we used machine learning to predict Binding energies of nuclei by training it on experimental values. We compare the predictions with the experimental values and see how accurate it is as compared to the theoretical model. We later suggest to apply our machine learning model on top of the liquid drop model to create a hybrid model.

Statement of the Problem:

Designing a nuclear model using Machine learning and comparing it with Liquid drop Model

Details of Method of Study:

Data:

The data is obtained from the Atomic Mass Data Centre (AMDC) website where the latest Atomic Mass Evaluation (AME2020) data is available in Fortran format. This data provides values of Atomic Mass, Mass Excess, Binding Energy, Beta Decay energy etc. We download this file and import the required columns and rows into our program for use.
Experimental Data: AME2020

Main Program:

We have divided the available data into two parts: Train set and Test set. This is done by randomly choosing a set amount of data points from the original data set. We call this our Training set and the remaining data points will be the Test set.

The train set will be used to train the machine learning algorithm and the test set will be used for comparing prediction accuracy of the models. The prediction accuracy is compared on the basis of rms error.

Liquid Drop Model (SEMF Model):

To get the Binding Energies of Nuclei we will use the Semi-Empirical Mass formula(SEMF).
The values of the five coefficients in the Semi Empirical mass formula have been empirically calculated by many researchers. Here we will be using values given by Rohlf J. W. in the book Modern Physics from α to Z^0, John Wiley & Sons.
Volume Energy term (aV)=15.75
Surface Energy term (aS)=17.8
Columb Energy term (aC)=0.711

Asymmetry Energy term (aA)=23.7
Pairing Energy term (delta)=11.18
The values of A, N and Z are imported from the data file. We define a function to calculate the Binding Energy using the Semi-Empirical Mass formula which takes N and Z as arguments.

Machine Learning Model:

In the Machine Learning model, we utilize the sklearn library of python to train the program using available Binding energy values. From the sklearn library we import the DecisionTreeRegressor tree function.

The function when called builds a regression model in the form of a tree by dividing the data into subsets and nodes. The depth of the tree can be specified while calling the function. Higher depth gives more accuracy for an errorless data sample. When using noisy data, the maximum depth needs to be optimized to exclude the noisy samples. The AMDC data is considered to be highly reliable and accurate and hence we do not need to account for noise in the program.

After calling the function we fit the curve to our data using .fit and then we predict the unknown values using the .predict function.

We train the Machine Learning algorithm on the experimental data by taking Atomic mass as a feature. (Independent Model)

We try to further improve the accuracy of the independent model by taking the predictions of SEMF model as a feature. We name this model ‘Level 0’.

We calculate the rms error in the predictions all models with respect to the experimental data and compare the values to determine the efficiency of each.

Further correction in the prediction of the Level 0 model can be achieved by creating an Error Model which predicts the error in prediction of Level 0 model.

The error here is the difference between the predictions made by the Level 0 model for the Training set and original the Training set data. We call this Error Training set and it is used to train the Error model to make predictions of the errors made by Level 0 model. This predicted error is added to the predictions of Level 0 model to compensate for the error. We call this Error Model ‘Level 1’.

Repeating the same procedure, we build another such error model to estimate the error of Level 1 model. We call this model ‘Level 2’.

From observations we can see the error decreases when the Error models are used. We stop building error models at Level 2 as observations show that higher level error models overcorrect the predictions and increase the rms error.

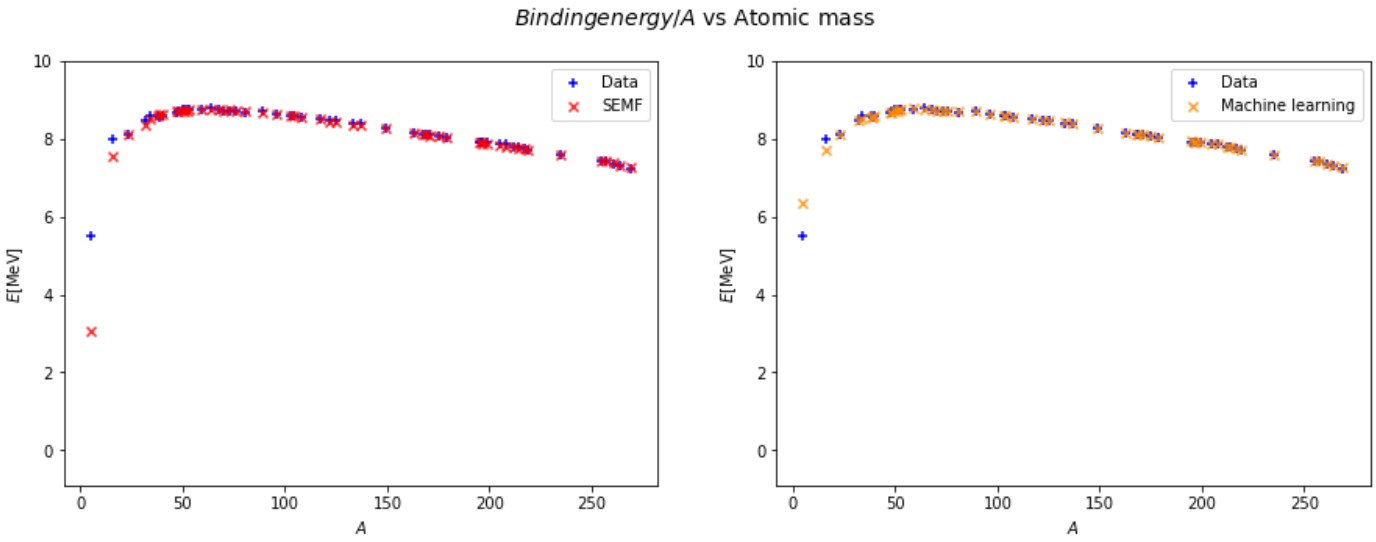
Results and Discussions:

We test our models for two different cases: Stable Isotopes and All Isotopes.

For stable isotopes we group our data by Atomic Mass number and assign it the value of maximum Binding Energy of the group.

Independent Machine Learning Model and SEMF

For stable isotopes

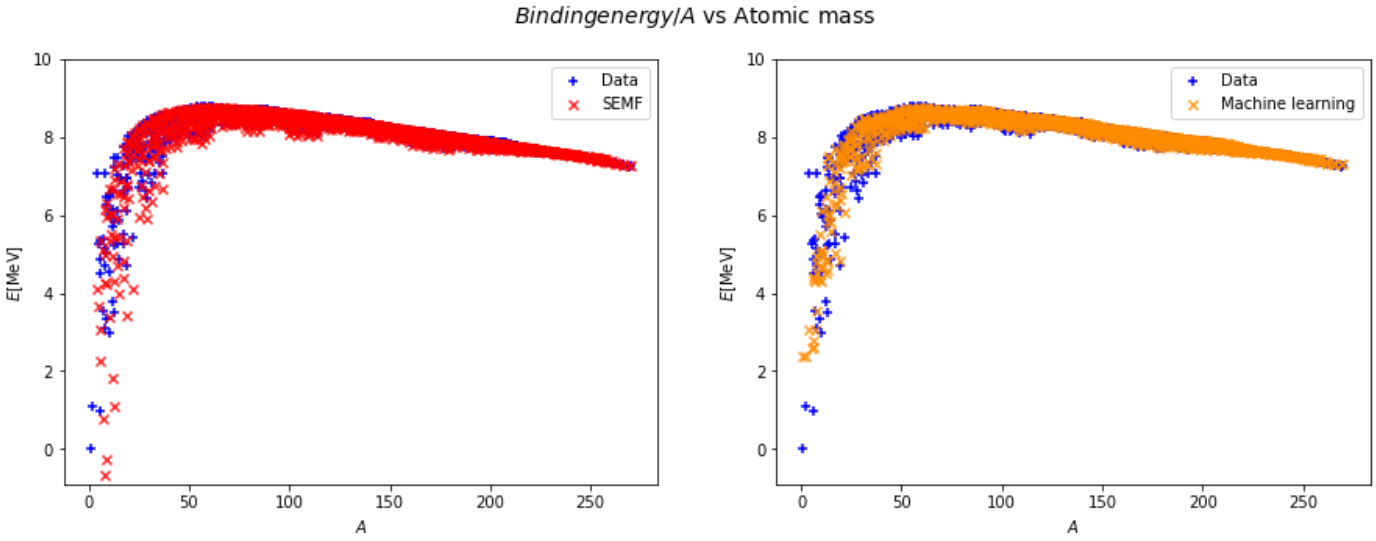


Error for Independent ML model: 0.122314543850655

Error for SEMF: 0.3400963895933706

SEMF model is known to give inaccurate values lighter nuclei which can be verified from the above plot. The Independent Machine learning model gives comparatively more accurate values, even for lighter nuclei.

For all isotopes



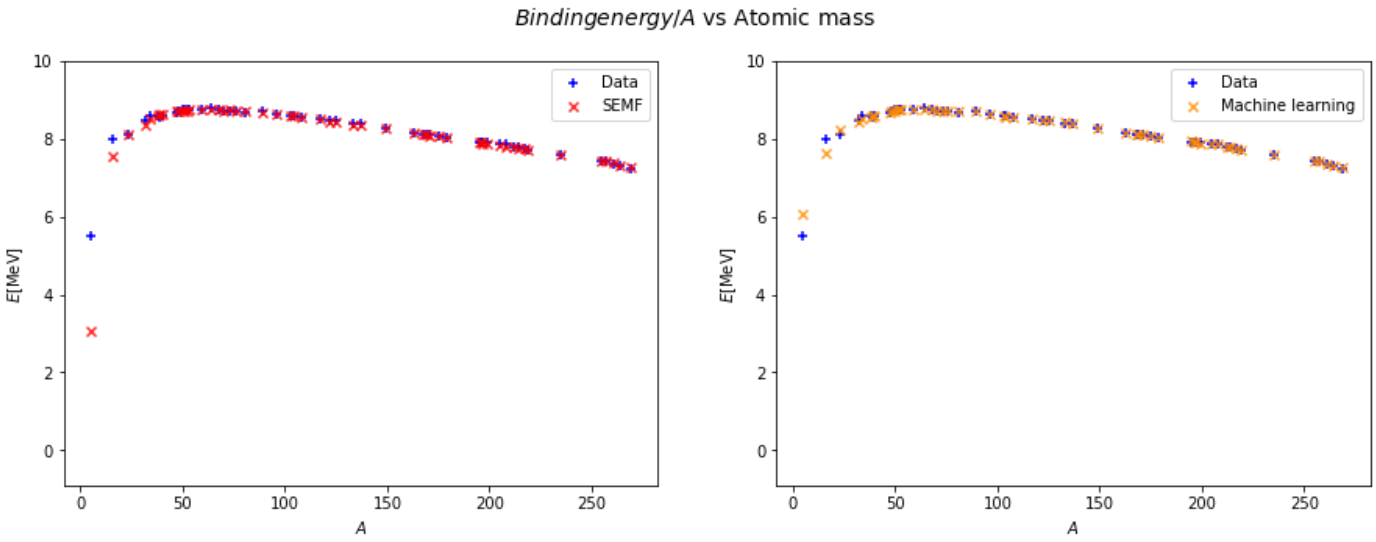
Error for Independent ML model: 0.21235006695813696

Error for SEMF: 1.0382852227640964

For all isotopes the prediction of Binding Energy becomes more difficult since there are multiple nuclei with the same atomic mass number. Hence our Independent ML model the error increases from ~0.1223 to ~0.2123. The Independent ML model still performs significantly better than the SEMF model. From the plots we can see that the prediction of the SEMF model deviates more from the experimental data than the Independent ML Model.

Hybrid Machine Learning Model And SEMF

For stable isotopes

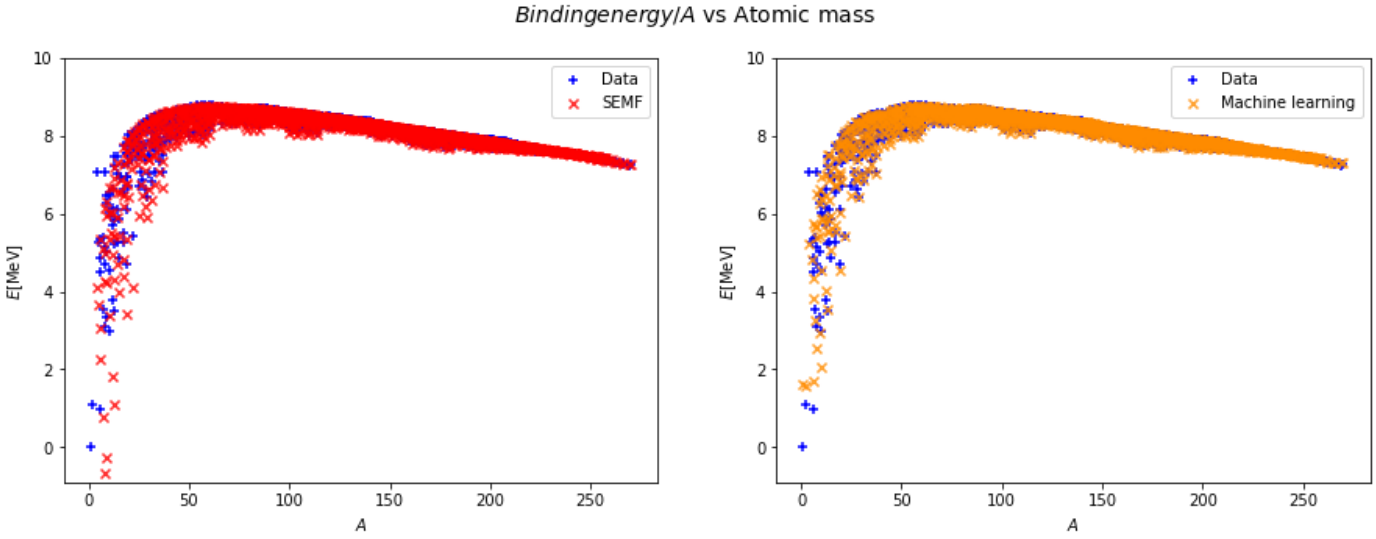


Error for Hybrid Machine Learning model: 0.09300376764642405

Error for SEMF: 0.3400963895933706

It is apparent from the plots that the Hybrid ML model performs better than the Independent ML model.

For all isotopes



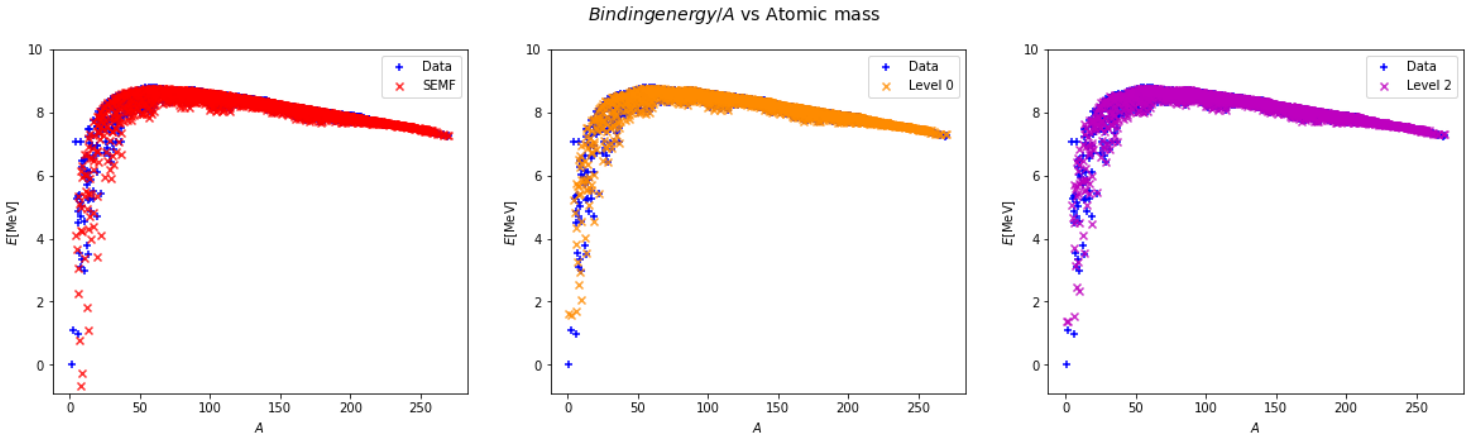
Error for Hybrid Machine Learning model: 0.09464530505887712

Error for SEMF: 1.0382852227640964

As expected, the Hybrid Model performs better than the SEMF and the Independent Model even for all the isotopes.

Error Model And SEMF:

The Error models try to improve the prediction accuracy of the Hybrid Model. (Level 0)



Error for level 0: 0.09464530505887712

Error for level 1: 0.09164359973398399

Error for level 2: 0.09055768069391838

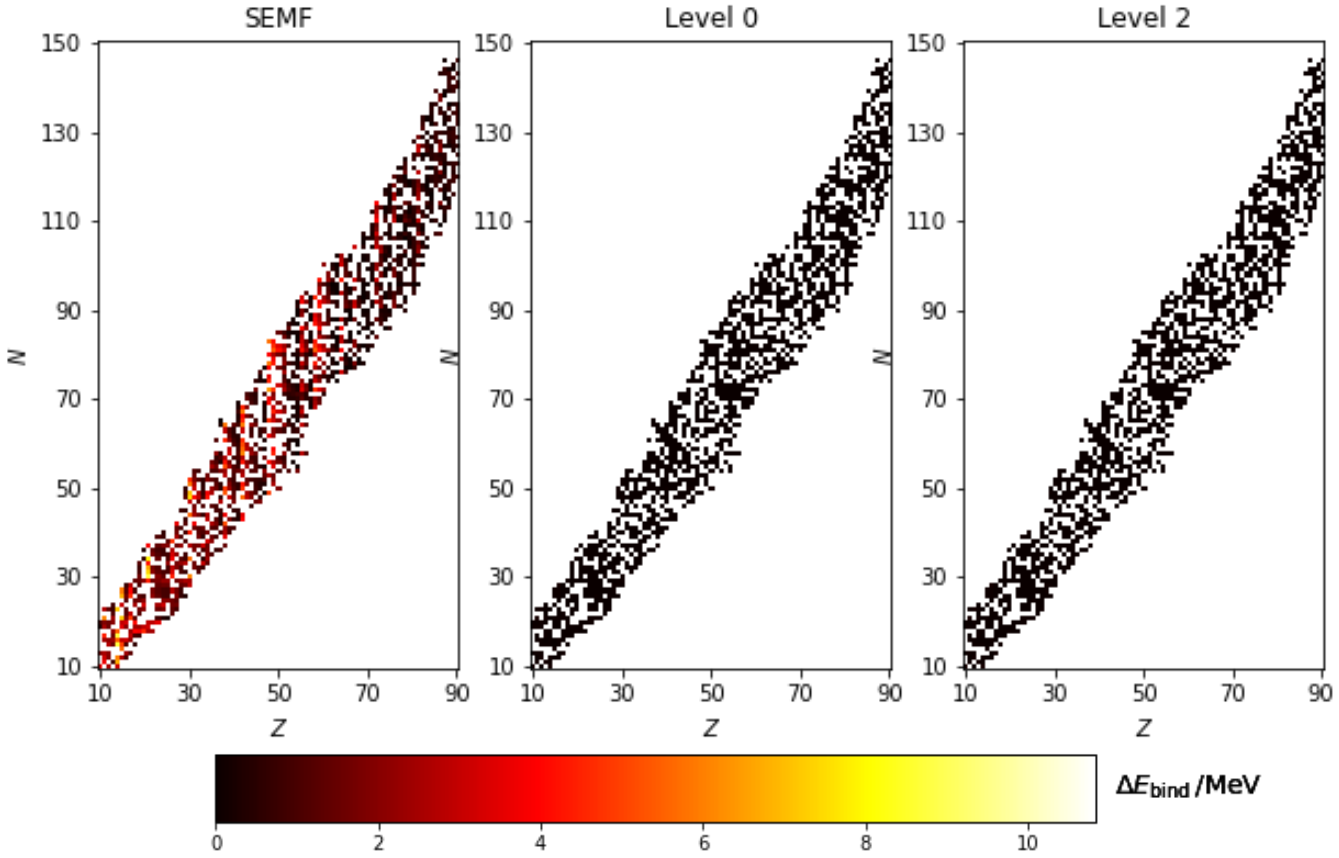
The results show that the error decreases with each layer. Level 2 is more accurate than Level 1 and Level 1 is more accurate than the Level 0 (i.e. the original Hybrid Model) Hence the error models successfully decrease the error in Hybrid ML model.

Following table shows Errors in each model for 10 random nuclei:

Error in SEMF	Error in level0	Error in level1	Error in level2
37.641	1.599988	1.407332	1.38999
3.72504	0.465793	0.273137	0.255795
2.985115	1.859107	1.996342	1.997141
1.622067	0.428442	0.556306	0.586139
5.605903	0.750883	0.688926	0.596522
1.818423	0.535095	0.42579	0.386172
0.035034	0.419624	0.356115	0.364391
2.209708	0.643626	0.767126	0.787699
0.25987	0.28114	0.228559	0.220671
2.79429	0.293024	0.38015	0.416467

For entire table: Click Here

Deviation Graphs:



The deviation of the predicted values from experimental values is indicated in this colour map. The darker shades represent less deviation We can see from the above plots that the SEMF model deviates significantly more from the experimental values than the level 2 and level 0 ML model. Hence we can say that the Level 2 model is more accurate than the SEMF model..

Limitations:

1. As the hybrid model is created on top of the Liquid Drop model, the problems of the theoretical model are also inherited by the hybrid model to some extent. This can be improved by using a better theoretical model the base model.
2. The data set for training our machine learning model is limited by experimental measurements of the new nuclei. Therefore, it becomes difficult to obtain more data for training such Machine Learning models.
3. Even though the predictions of our Machine learning models were better than the Liquid drop model, there is still a substantial error in the predictions for lighter nuclei.

Relevance of the Study:

1. Machine Learning Models built on top of theoretical base models perform better than both independent machine learning models and theoretical models. In some experiments it might be more useful than the theoretical models.
2. Some nuclei occurring in astrophysical systems cannot be experimentally studied in laboratories. Measurements of Binding Energy of these nuclei is a hot topic in nuclear physics. Hence accurate nuclear models become very necessary to predict the binding energy of such unknown nuclei. We can use machine learning models to refine the inaccuracies of the theoretical models and hence give more accurate results. They are suitable for extrapolation and interpolation for data points close to the original set.

3. Development of more efficient machine learning models can help reveal more physical hints from predictions and help us improve present nuclear models.

Bibliography:

1. Introductory Nuclear Physics (Book by Kenneth S. Krane)
2. Nuclear Talent course on Machine Learning in Nuclear Experiment and Theory
3. Nuclear Physics (by D.C. Tayal)
4. Modern Physics from α to Z^0 (Book by Rohlf J. W)
5. Tianliang Zhao, Hongfei Zhang, A new method to improve the generalization ability of neural networks: A case study of nuclear mass training, Nuclear Physics A, Volume 1021, 2022, 122420, ISSN 0375-9474, https://doi.org/10.1016/j.nuclphysa.2022.122420.
6. Gao, Zepeng & Wang, Yongjia & Lü, Hongliang & Li, Qingfeng & Shen, Caiwan & Liu, Ling. (2021). Machine learning the nuclear mass.
7. Nuclear mass predictions for the crustal composition of neutron stars: A Bayesian neural network approach, R. Utama, J. Piekarewicz, and H. B. Prosper, Phys. Rev. C 93, 014311 – Published 20 January 2016

Python Programme:

CODE LINK