Nuclide Identification in Gamma-ray Spectroscopy Using Mixture of Experts Approach

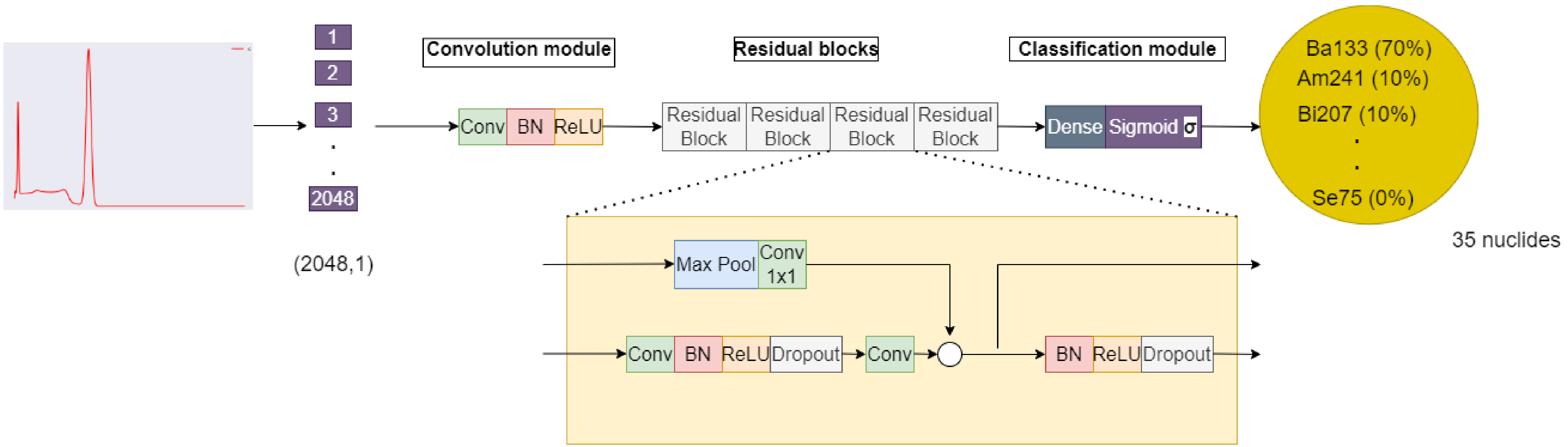
Abdulkader Helwan, Mats Eriksson, Alexandr Malusek

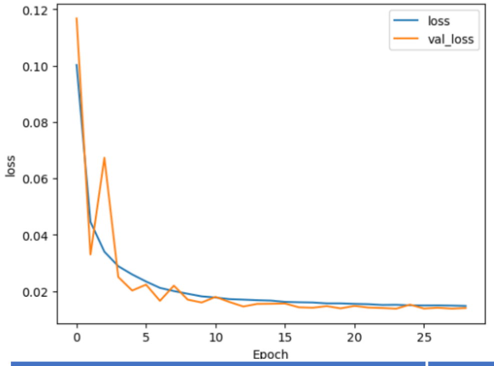
LINKÖPING UNIVERSITY

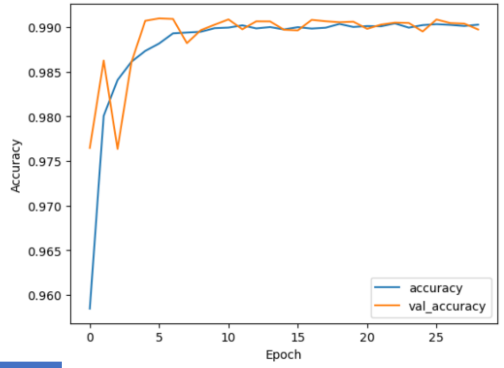
Objectives

- Gamma Ray Radionculdes Identification in Spectra
  - Activity prediction
  - Nuclides identitfication
- Develop a global model with less constraints in terms of number of nuclides identified and method of inputting the spectrum.
- Create a large-scale dataset of spectra

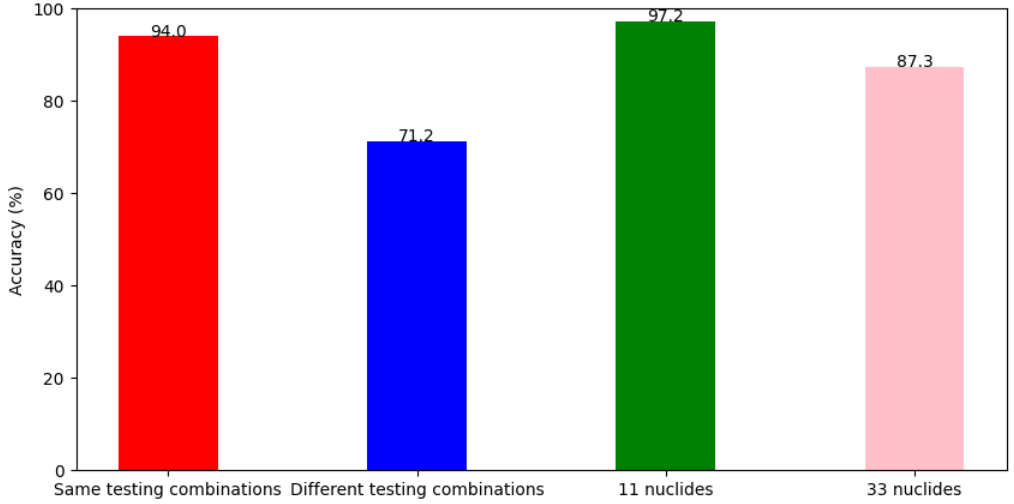
Results







(b)

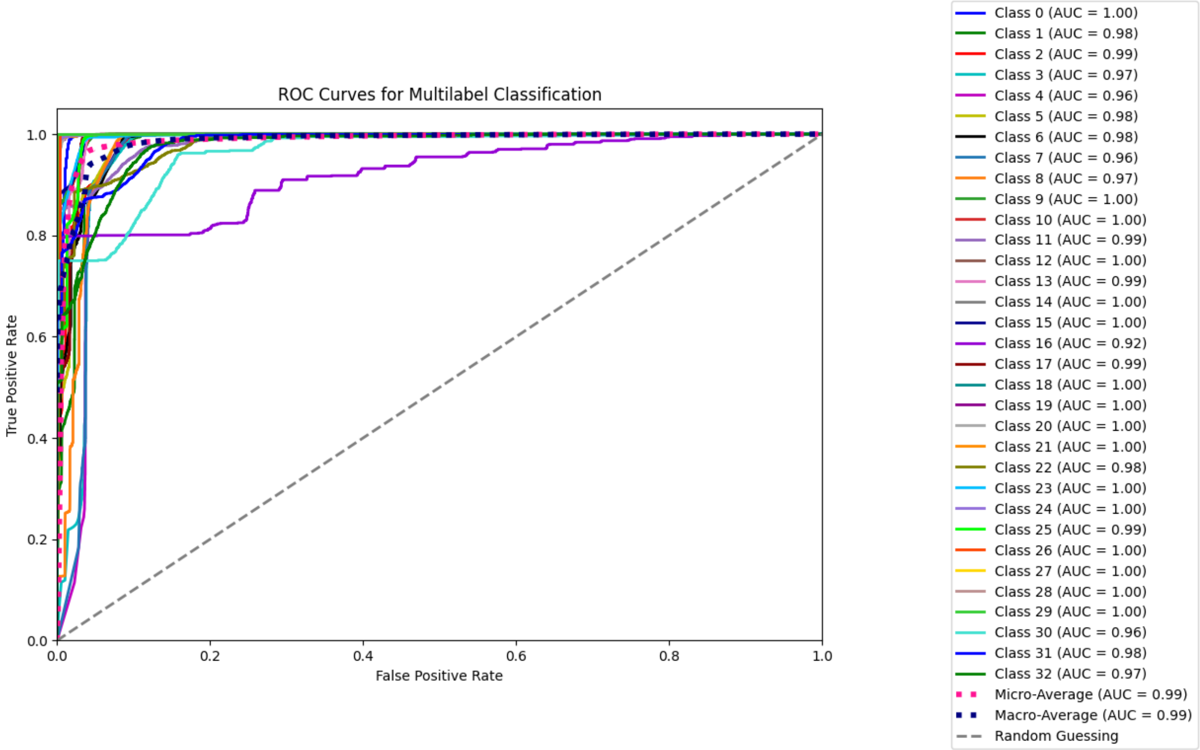


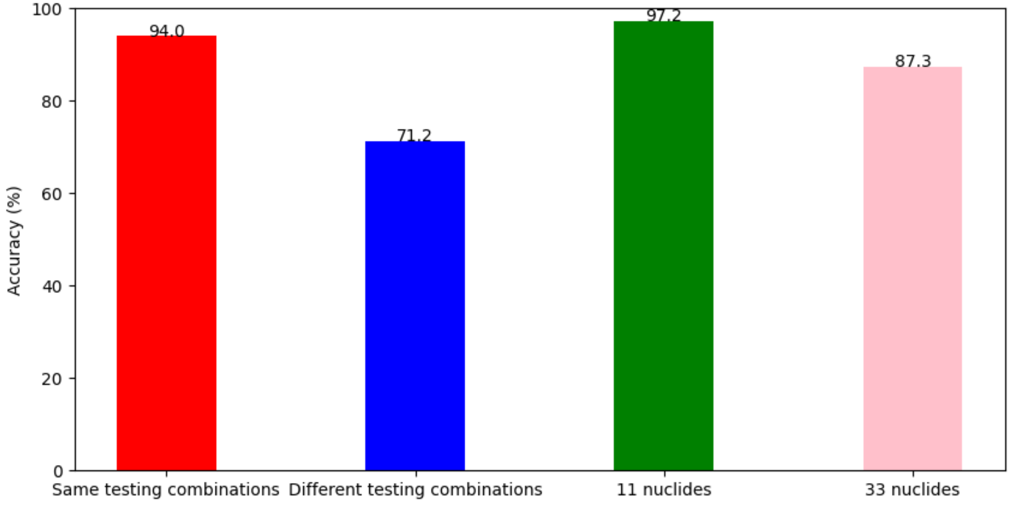
Testing evaluation metrics	Values (%)
Accuracy	0.94
Hamming loss	0.011
F1-score	0.966
Precision	0.965

	Accuracy	Hamming loss
ANN [8]	76.61%	0.239
CNN [27]	93.2%	0.029
Ours (ResNet)	94%	0.011

Limitations

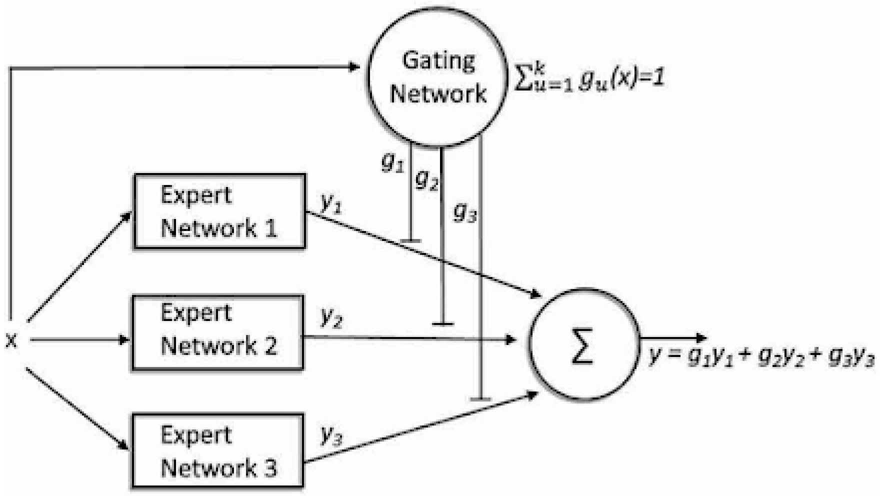
- Same combinations
- Large number of nuclides
- Overfitting





Proposed Method

- The MoE model is a type of ensemble learning method that combines the predictions of multiple “expert” models, each of which specializes in a specific portion of the data.
- A gating function, or router, decides which expert(s) should handle each input data point.
- Both the experts and the router are trained simultaneously using gradient descent, a common optimization algorithm in machine learning.

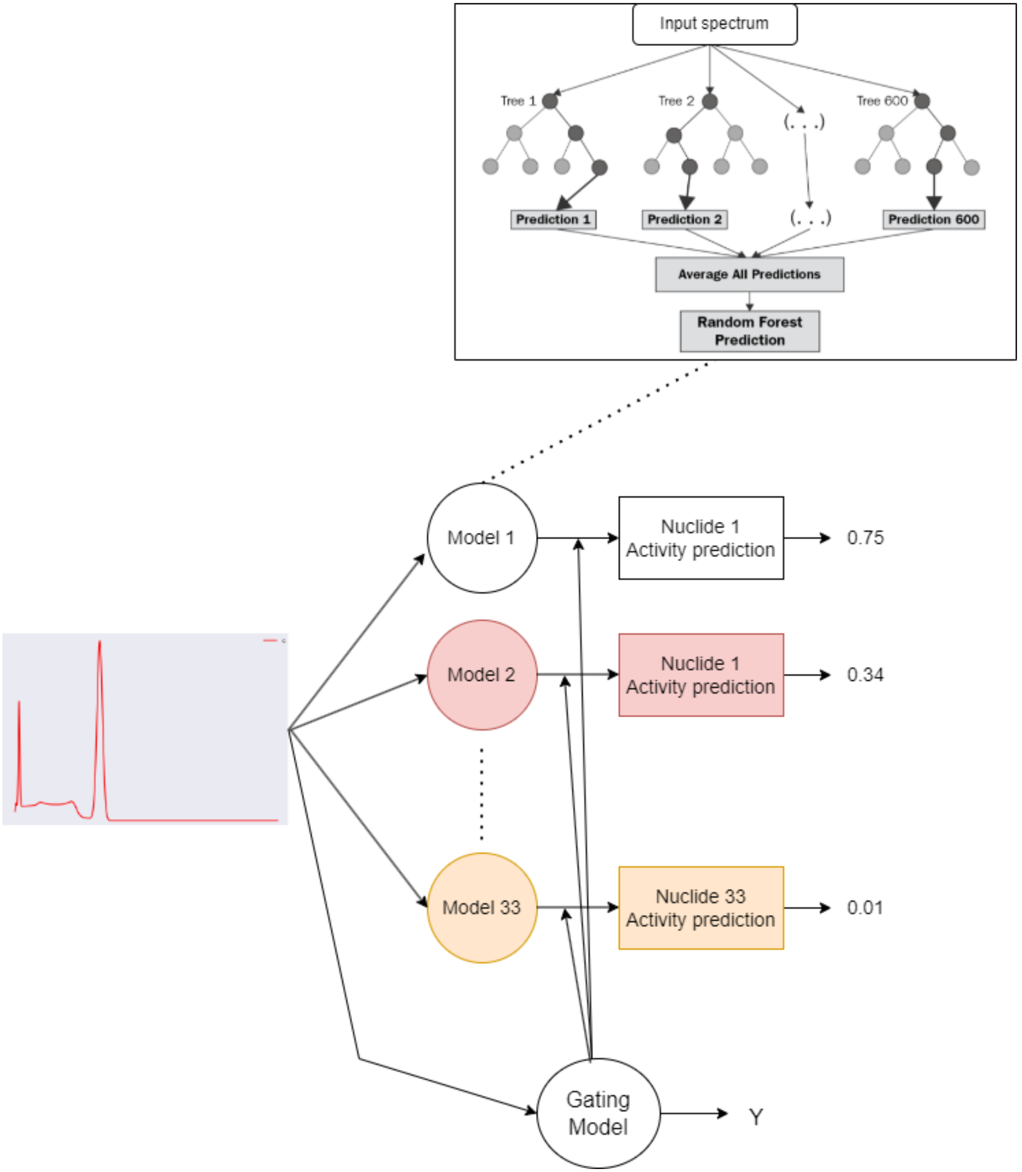


https://arxiv.org/abs/2208.02813

Mixture of Experts

In the MoE model, we have 33 experts, each specializing in recognizing a specific one nuclide from 1 to 33.

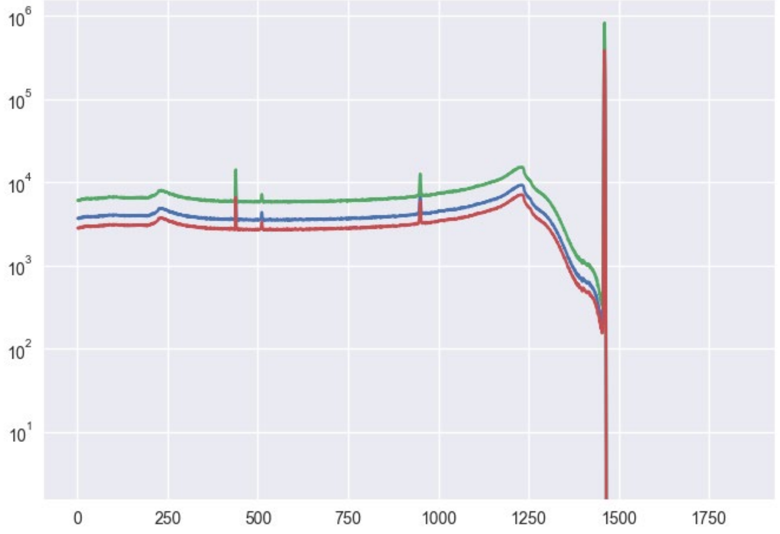
- Each expert is Random Forest that has been trained to recognize its specific nuclide.
- Alongside the experts, there's a router, another neural network that takes a spectrum as input and decides which expert should handle it.
- For example, if the router receives a spectrum of nuclide '3', it should route this spectrum to the expert that specializes in recognizing nuclide '3's.
- During training, both the experts and the router are trained using gradient descent. The experts are trained to become better at recognizing their specific nuclides, and the router is trained to become better at dispatching spectra to the right expert.
- The output of the MoE layer is a combination of the outputs of the selected experts, weighted by the router's gating values.



Dataset

Isotope	Half life
Xe-137	3.818 m
Y-95	10.3 m
Ba-142	10.6 m
XE-138	14.08 m
LA-143	14.2 m
MO-101	14.61 m
Y-94	18.7 m
BA-141	18.27 m
TE-133	12.5 m
TE-134	41.8 m
SB-131	23.03 m
TC-104	18.3 m
TE-133M	55.4 m
MN-56	2.57 h
SR-92	2.71 h
KR-87	76.3 m
KR-88	2.84 h
I-135	6.57 h
CS-138	33.41 m
I-134	52.5 m
XE-135M	15.29 m
Y-93	10.18 h
SR-91	9.63 h
SN-128	59.07 m
SB-130	39.5 m
ZR-97	16.91 h
KR-85M	4.48 h
SB-129	4.40 h
TE-131	25 m
MO-99	65.9 h
TE-132	3.2 d
CO-58M	9.04 h
BA-140	12.75 d
I-133	20.8 h
LA-142	91.1 m

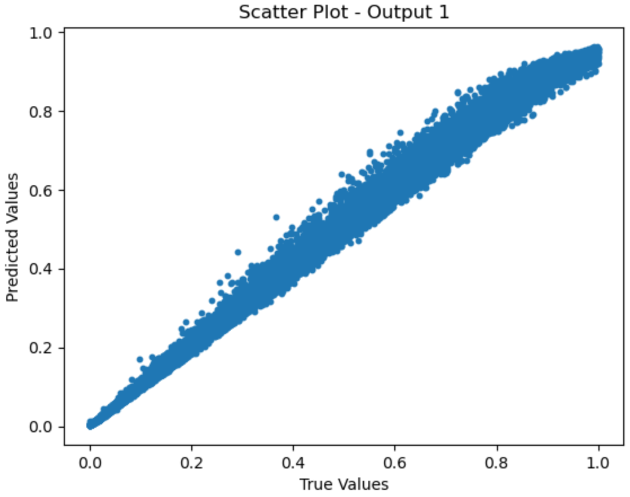
- The gamma-ray spectrum S_l of lth radionuclide ($1 \leq l \leq 33$) was calculated using the Nucleonica Gamma Spectrum Generator Pro application.
- The setup was: a high-purity germanium detector (HPGe) with an input window of Beryllium, a crystal packing with a crystal diameter of 80 mm and height of 30 mm, and a reflector of germanium.
- The source was placed at the detector axis 5 mm from the entrance window. The number of spectrum channels was 2048. The channel-to-energy conversion factor was set to 0.9 keV/channel, and the full width at half maximum (FWHM) was 0.75 keV at 122 keV. The nuclide activity was set to 1 Bq, and the acquisition time was set to 1000 s.

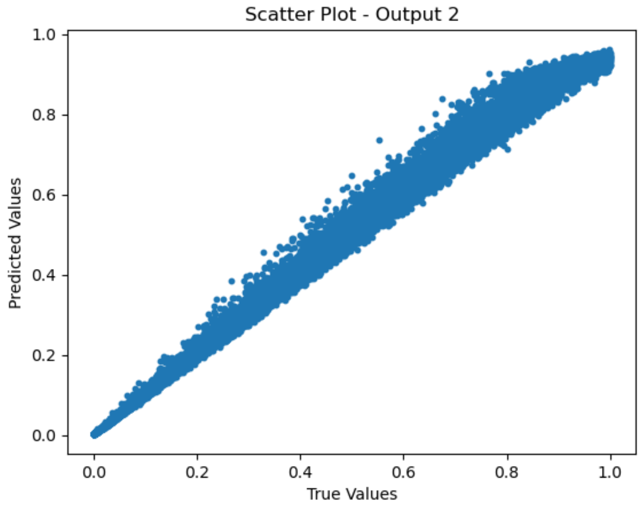


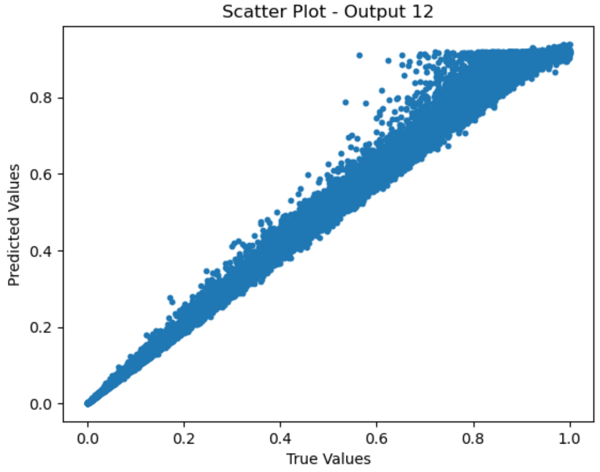
An example spectrum of ^{40}K. (Blue) shows the original spectrum, (b) and (c) show the same spectrum with different random activity coefficients.

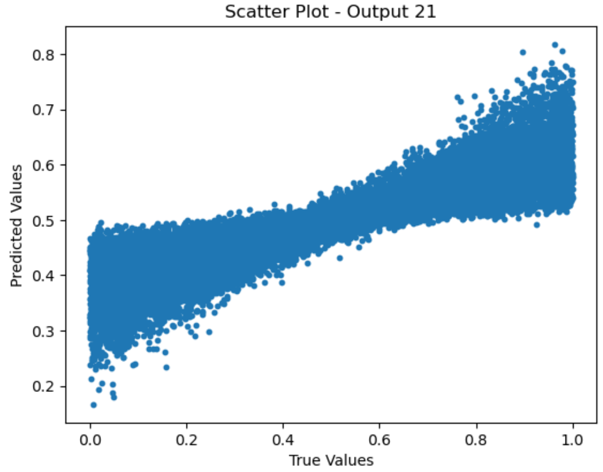
Results

Nuclides	Root Mean Squared Error (RMSE)	Correlation Coefficient
1	0.023703013623059353	0.9967496885208839
2	0.025121475463594677	0.9963192561813509
3	0.024476372428756468	0.9964944641969723
4	0.023702739855678043	0.9966920050692459
5	0.02659358946776954	0.9958850399225772
6	0.024780980561838407	0.9964158170540809
7	0.024134627599482036	0.9965832269255369
8	0.03088518930724672	0.9943166795114441
9	0.026278036629398992	0.9959401506855421
10	0.02415947683798635	0.9965852143067994
11	0.03118459850214415	0.9942510391835282
12	0.0291896508609519	0.9949522171829567
13	**0.23335315644982754**	**0.8268628819317632**
14	**0.2179365736822283**	**0.9016923034249824**
15	**0.2141924010002705**	**0.8962067227615111**
16	**0.23207277845387986**	**0.8308446717258501**
17	**0.2089857451160745**	**0.9284366694862533**
18	**0.218503800059807**	**0.8937213884201568**
19	**0.21385891619668396**	**0.9135922768156411**
20	**0.2240794627866336**	**0.8686029093256844**
21	**0.21840890955649808**	**0.8915431135062085**
22	**0.22512363128803964**	**0.8348881403122959**
23	**0.2241679888290192**	**0.8799150174994943**
24	**0.22192440172407**	**0.8984664513684325**
25	**0.21332151853666442**	**0.9192021956105032**
26	**0.22676775971211405**	**0.8636938228600154**
27	**0.21592741167770205**	**0.9004007269387128**
28	**0.2266491084196206**	**0.8422502437599863**
29	**0.2246842226808011**	**0.8715486299479897**
30	**0.23288800132255452**	**0.8432124303767998**
31	**0.2143519277411918**	**0.9183255316604655**
32	**0.22481984259297416**	**0.8841746143105487**
33	**0.20856686697244903**	**0.9237930858598694**



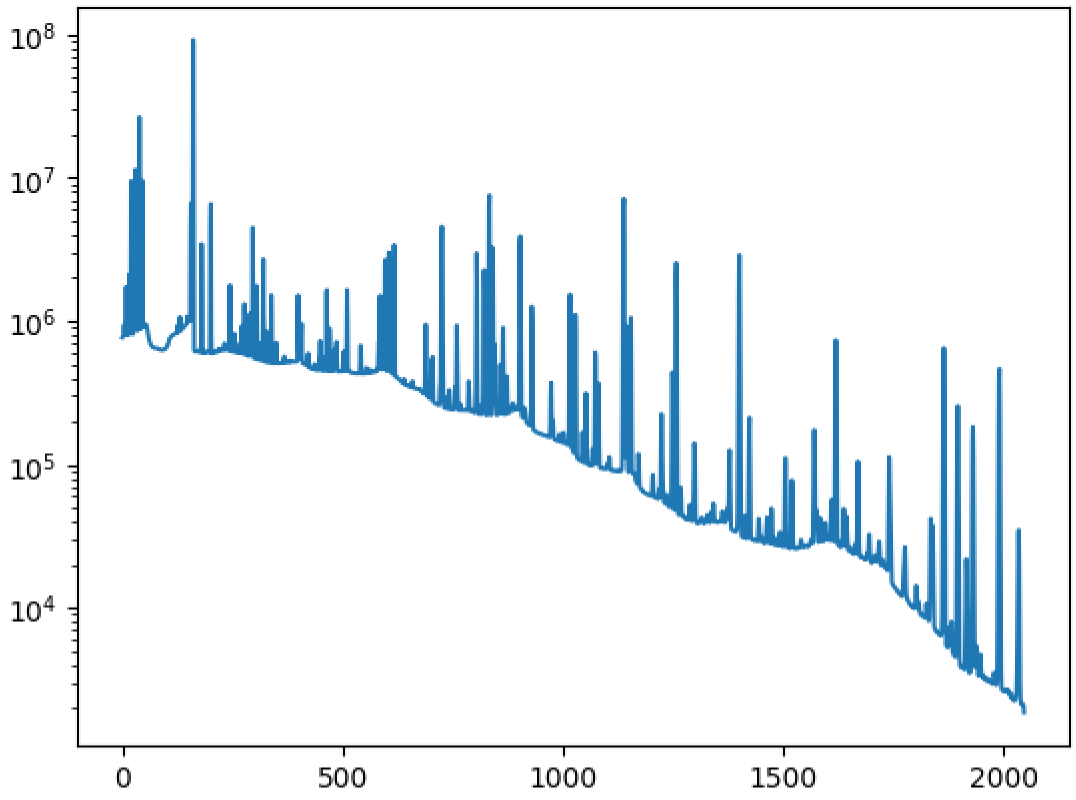






Dataset 2: Fine-tuning or Retraining the MoE

	Nuclides
1	Y-93 - Yttrium (Y)
2	I-135 - Iodine (I)
3	Sr-91 - Strontium (Sr)
4	Mo-99 - Molybdenum (Mo)
5	La-141 - Lanthanum (La)
6	Ce-143 - Cerium (Ce)
7	Sb-129 - Antimony (Sb)
8	Te-129 - Tellurium (Te)
9	Ba-140 - Barium (Ba)
10	Zr-95 - Zirconium (Zr)
11	Ru-103 - Ruthenium (Ru)
12	Ce-141 - Cerium (Ce)
13	Sr-89 - Strontium (Sr)
14	Y-91 - Yttrium (Y)
15	Ce-144 - Cerium (Ce)
16	Pr-143 - Praseodymium (Pr)
17	I-131 - Iodine (I)
18	Sb-128 - Antimony (Sb)
19	Te-131 - Tellurium (Te)
20	Te-132 - Tellurium (Te)

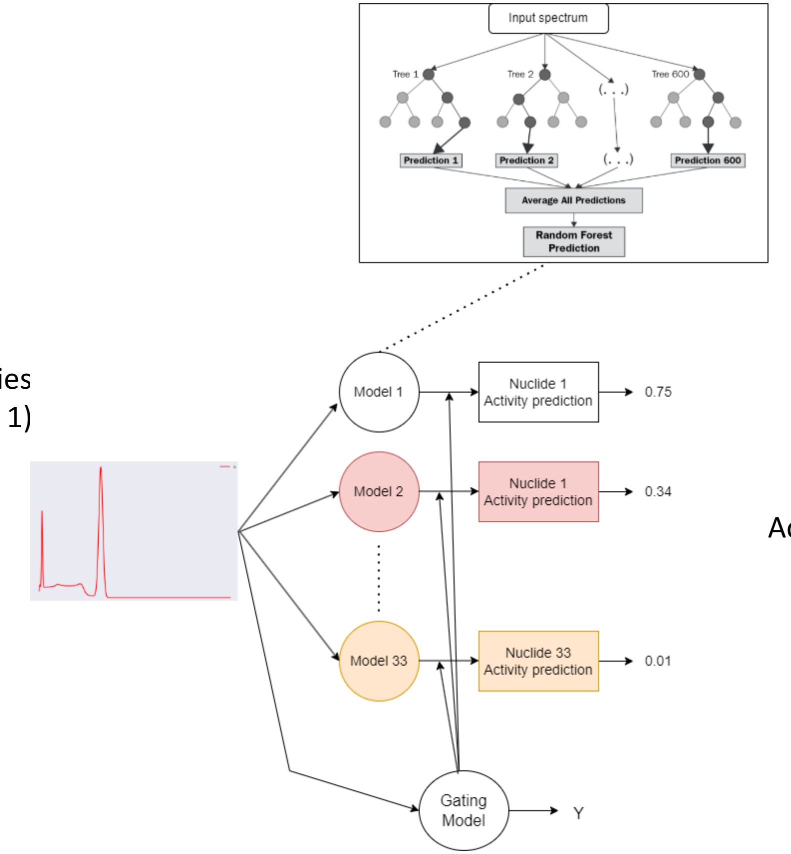


Combined Model Mean Squared Error: 0.0537

Training on 11 Nuclides

	Nuclides
1	Y-93 - Yttrium (Y)
2	I-135 - Iodine (I)
3	Sr-91 - Strontium (Sr)
4	Mo-99 - Molybdenum (Mo)
5	La-141 - Lanthanum (La)
6	Ce-141 - Cerium (Ce)
7	Sb-129 - Antimony (Sb)
8	Te-129 - Tellurium (Te)
9	Ba-140 - Barium (Ba)
10	Zr-95 - Zirconium (Zr)
11	Sr-89

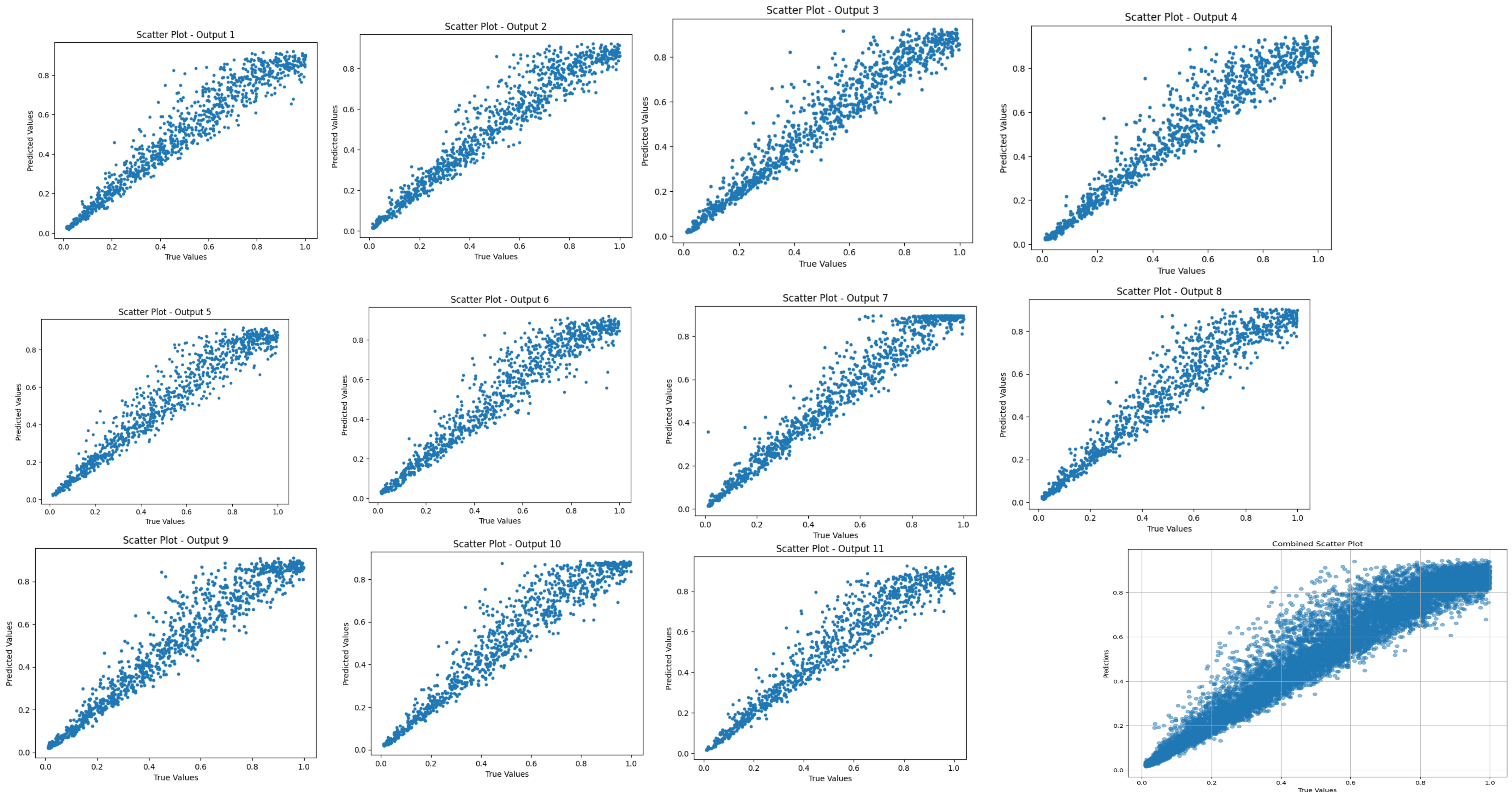
5000 examples
(Random activities
generation (0 to 1)



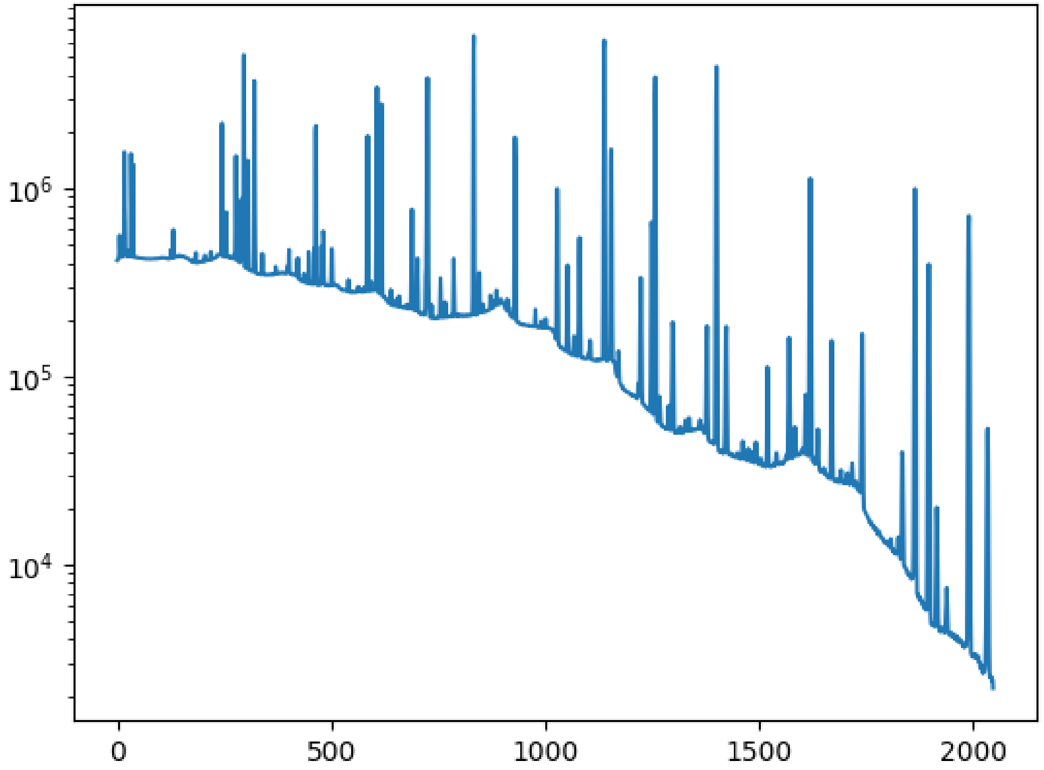
Activities predicition

Nuclides	RMSE	R 2
1	0.07	0.92
2	0.07	0.93
3	0.07	0.92
4	0.07	0.92
5	0.07	0.92
6	0.06	0.94
7	0.07	0.95
8	0.07	0.92
9	0.07	0.92
10	0.07	0.92
11	0.07	0.92
Average	0.058	0.9287

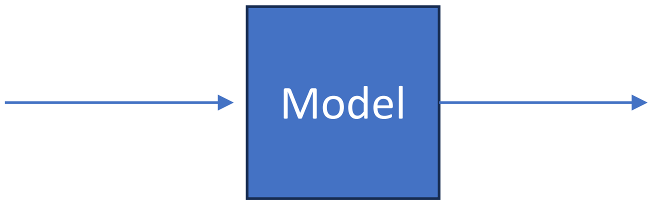
Testing Results



Inference

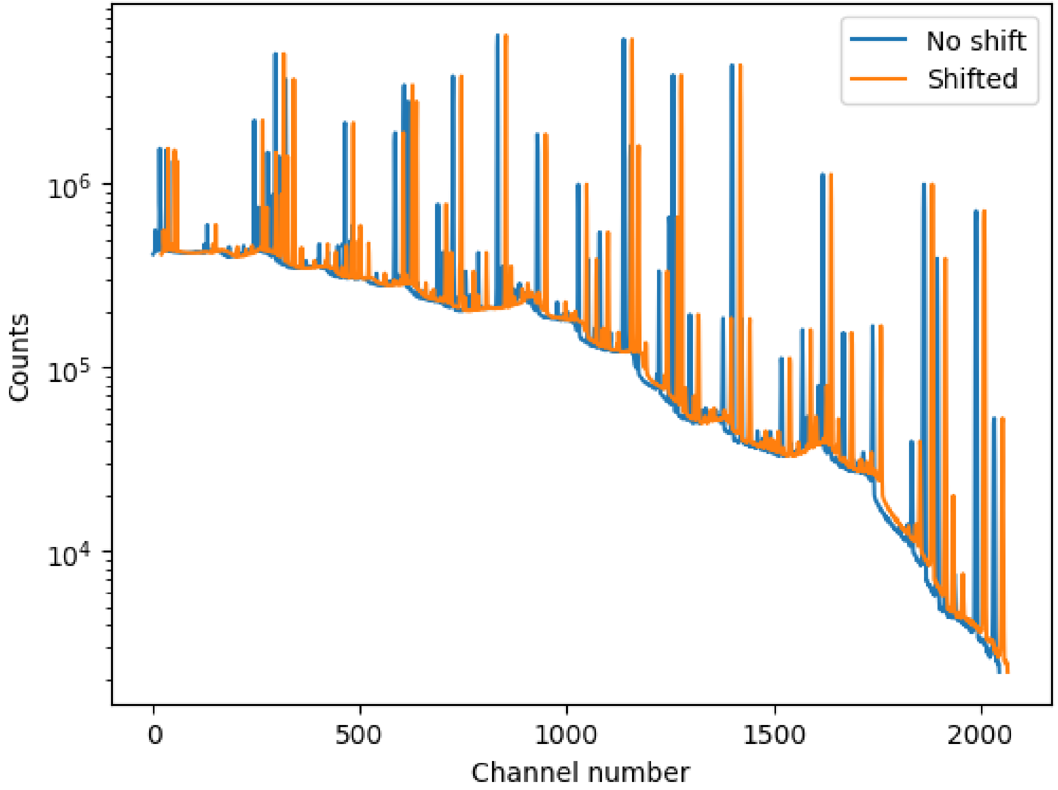


Nuclides: Y-93, I-135, Sr-91
Activities: 0.46, 0.87, 0.83

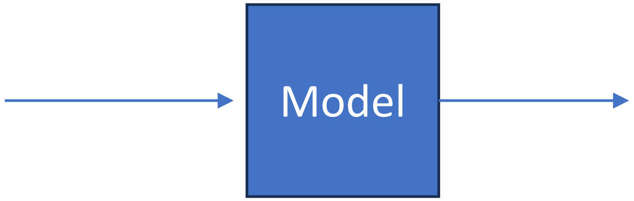



```
Expert Predictions:
Nuclide 1: 0.8543949999999991
Nuclide 2: 0.7937049999999997
Nuclide 3: 0.7971250000000002
Nuclide 4: 0.09927000000000001
Nuclide 5: 0.28562999999999966
Nuclide 6: 0.014739999999999974
Nuclide 7: 0.02434999999999997
Nuclide 8: 0.029294999999999963
Nuclide 9: 0.03645499999999996
Nuclide 10: 0.04495499999999999
Nuclide 11: 0.023644999999999958
```

Gain Shift



Nuclides: Y-93, I-135, Sr-91
Activities: Activities: 0.46, 0.87, 0.83
Gain Factor=20 channels




```
Expert Predictions:
Nuclide 1: 0.8543949999999991
Nuclide 2: 0.7937049999999997
Nuclide 3: 0.7971250000000002
Nuclide 4: 0.09927000000000001
Nuclide 5: 0.28562999999999966
Nuclide 6: 0.014739999999999974
Nuclide 7: 0.02434999999999997
Nuclide 8: 0.029294999999999963
Nuclide 9: 0.03645499999999996
Nuclide 10: 0.04495499999999999
Nuclide 11: 0.023644999999999958
```

Contributions and Future Works

- **Mixture of Experts Approach:** To address limitations and challenges of CNN-based models, we proposed a new approach called "Mixture of Experts."
- This approach allowed the model to be tested on spectra with different numbers of nuclides and activities, enhancing its capabilities in both nuclide identification and activity prediction.
- We are currently testing our second approach that overcomes the limitations and challenges faced in the previous one (deep CNN) on a new set of more complex nuclides with shorter half-life.
- This new method is meant to be a foundation model that can be fine-tuned to any set of nuclides in a short time, which makes it practical and globally used in several situations such as Nuclear Medicine, Radiation Therapy, Emergency Preparedness, Nuclear Power etc…
- More testing is needed
- More data is needed to increase the performance of the models