Jordan Ellis
School of Physical and Chemical Sciences
Queen Mary University of London,
Supervised by Adrian Bevan

Radioisotope identification using machine learning.

Abstract:

Identifying radioactive isotopes requires different forms of radiation spectroscopy, this would require access to a lab environment. To help tackle this problem, I wrote a program in Python using TensorFlow to do machine learning to understand how different radioactive nuclei decay. The algorithm made predictions and differentiated between radioisotopes. It can be further developed to detect radioactive isotopes in the real world. However, when radioactive isotopes have long half-lives, the algorithm is unable to differentiate between them.

Bateman Equation:

The Bateman equation is a mathematical model describing abundances and activities in a decay chain as a function of time, based on the decay rates and initial abundances.

$$N_n(t) = N_1(0) \times \prod_{i=1}^{n-1} \lambda_i \times \sum_{i=1}^{n} \frac{e^{-\lambda_i t}}{\prod_{j=1, j\neq i}^{n} (\lambda_j - \lambda_i)} \quad (1)$$

Where $N_n(t)$ = activity/count at an instance in time, $N_1(0)$ = the initial activity/count of the parent nucleus, λ_i = the decay constant the radioisotope and it decays into λ_{i+1} .

Simulating Data:

- Takes the half-life of each selected isotopes and creates a time series of the activity (up to 10 years in the example) using the Bateman equation.
- The program can separate each type of decay for example "Alpha", "Beta -" , "Beta +" and "Neutron emission" will be separated and they will have their own unique time series. This assumes a perfect detector.
- All categorical data such as labels for names of isotopes and types of decay are replaced with numerical values.

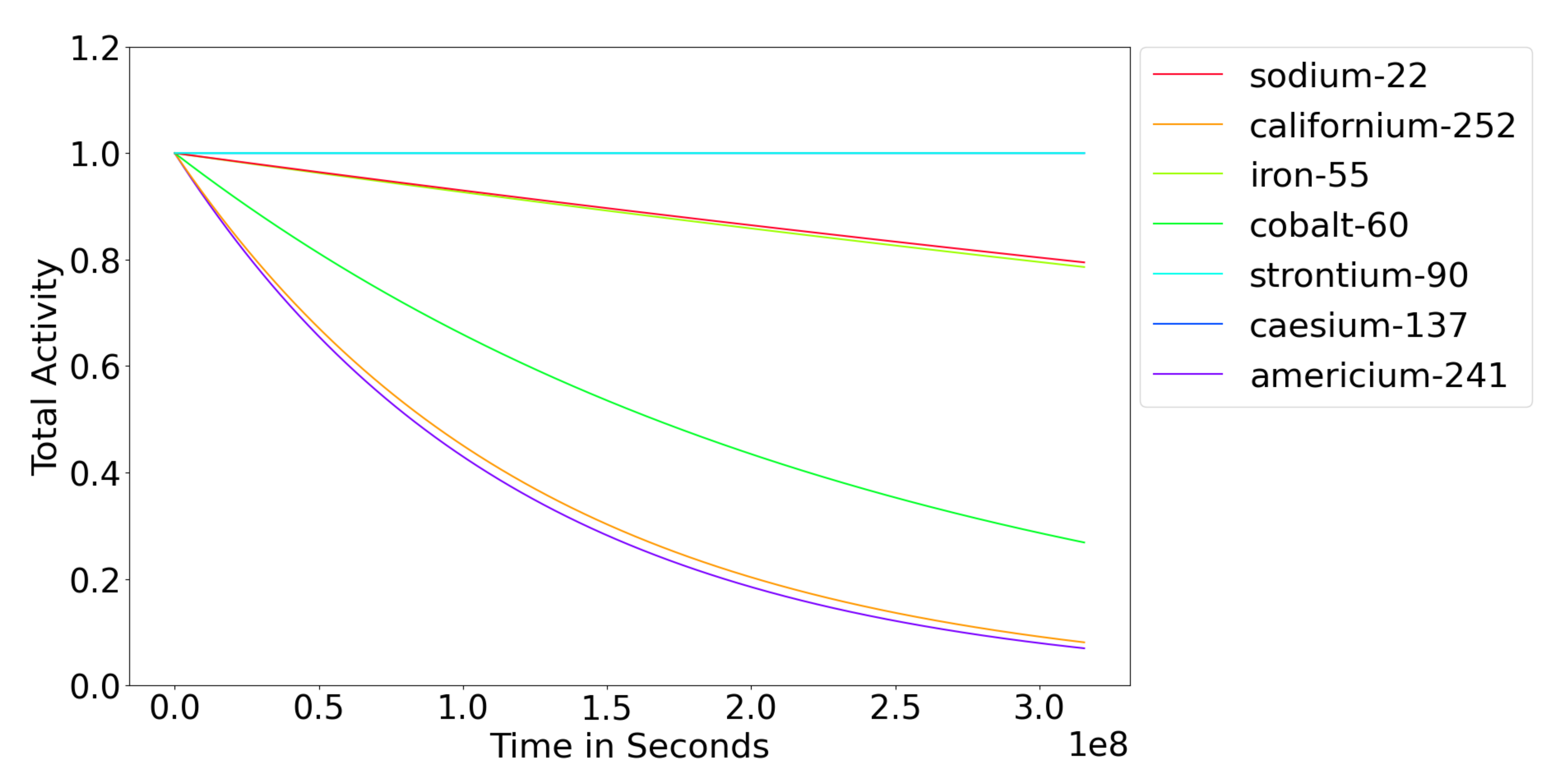


Figure 1 : Activity time series for 7 different radioisotopes that are used in a lab environment over 10 years.

Training & Evaluating:

- All activities and decay types are normalized between 0 and 1 to make it easier to train.
- The data is passed through a neural network with two hidden layers the first layer contains 16 nodes and the second layer contains 32 nodes. The output layer depends on the number of isotopes the user wants to identify.
- The models can be trained on any number of epochs, for lab environment example the model was trained on 100 epochs (see in Fig 2).
- To understand where the algorithm struggles, users can plot a confusion matrix shows which shows what the algorithm predicted versus against that the actual radioisotope.

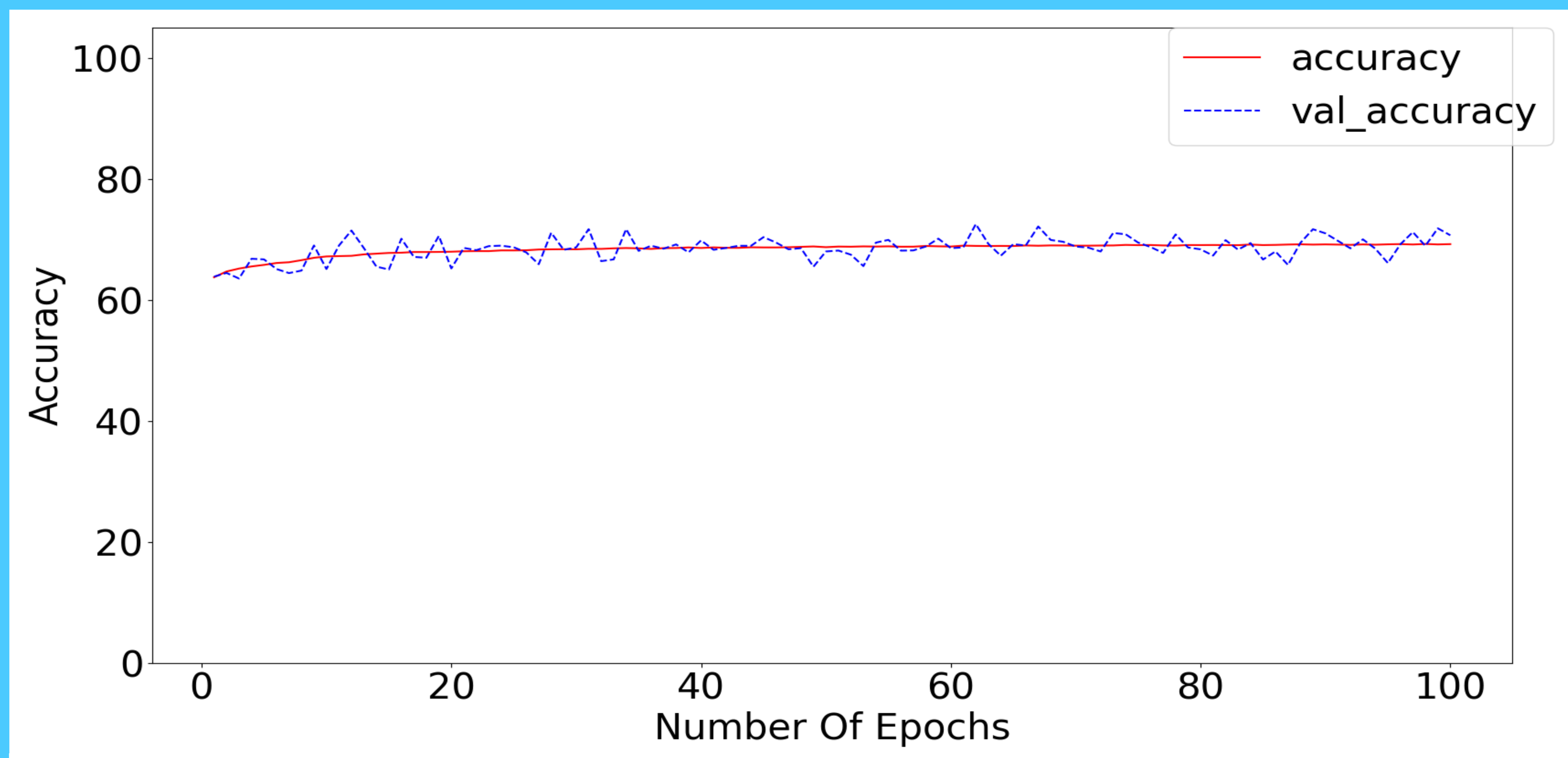


Figure 2 : Accuracy of model when training on data from isotopes in Figure 1 over 100 epochs.

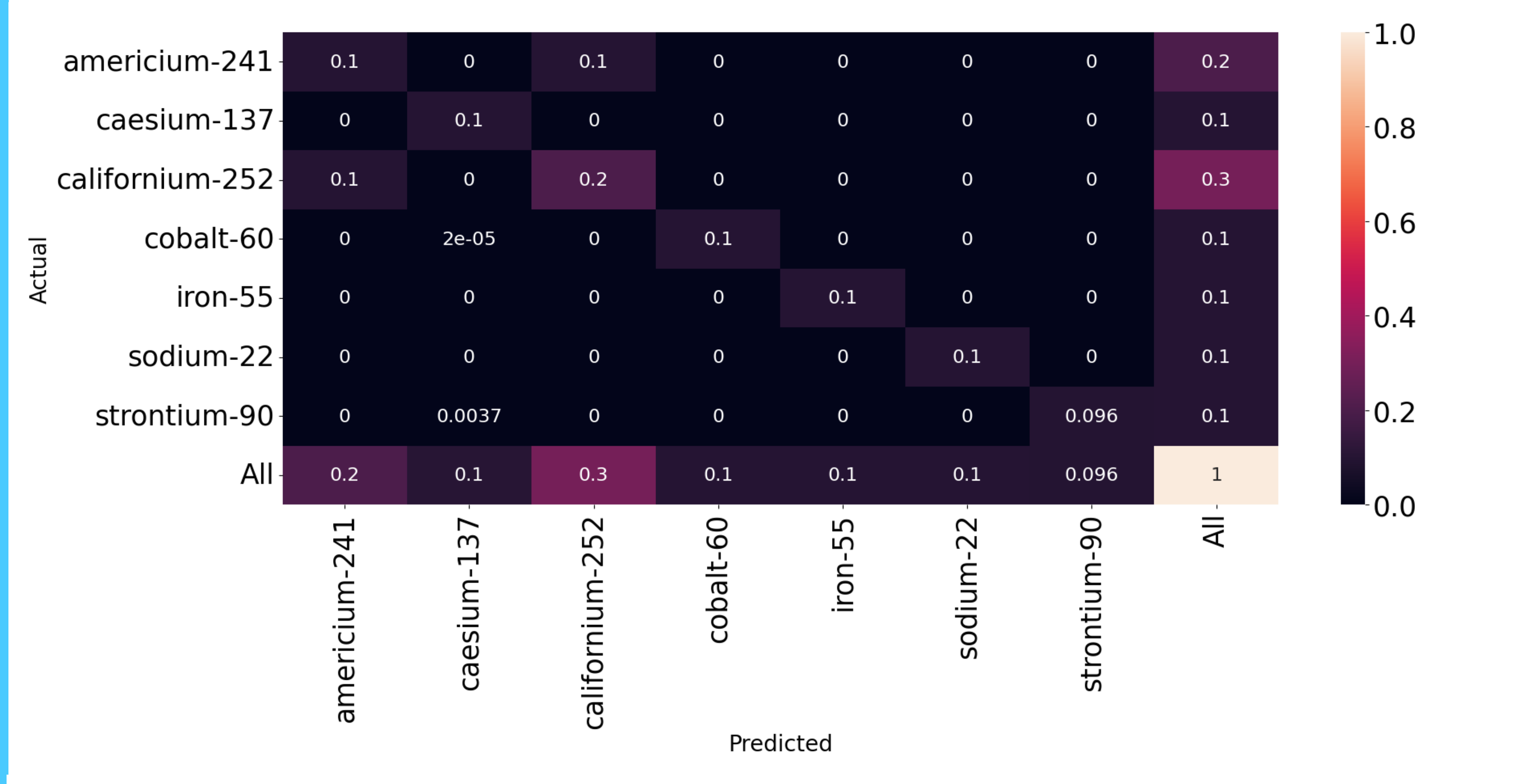


Figure 3 : confusion matrix for the trained model in Figure 2.

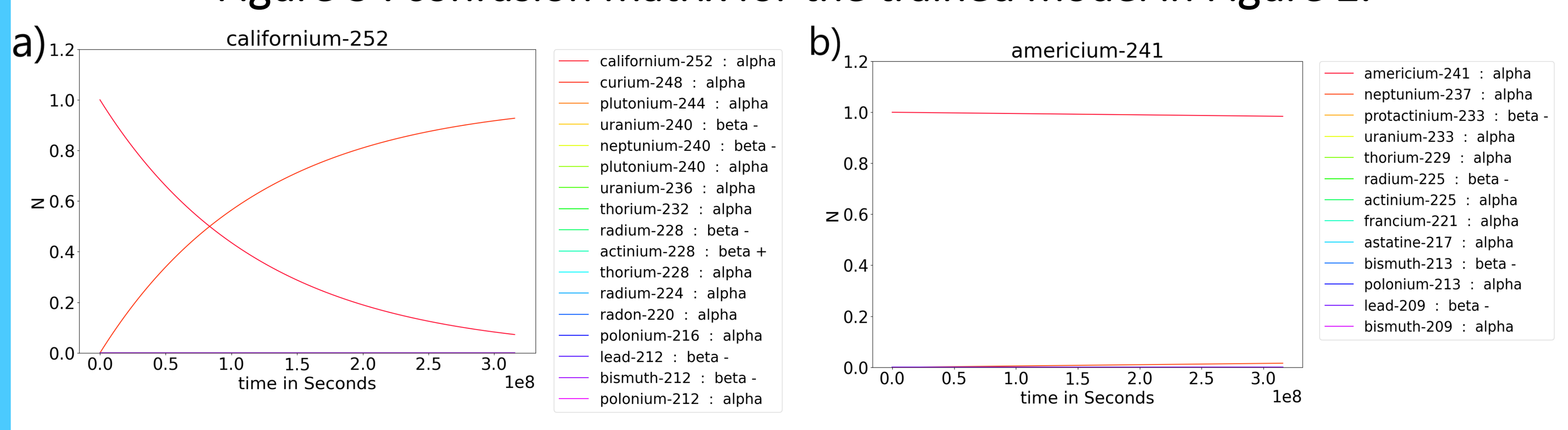


Figure 4 : a) Decay of californium-252 and all the other radioactive isotopes in its decay chain, only data for californium-252 and curium-248 are visible in the 10 years. b) Decay of americium-241 and all of the isotopes in its decay chain. Only americium is visible due to its long half-life.

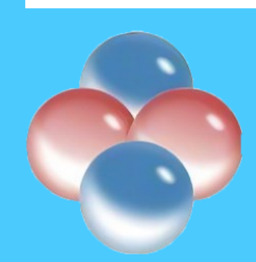
Results:

- Better than 65% accuracy as shown in Fig 1.
- Confusion matrix (Fig 2) shows some isotopes are almost always correctly Identified.
- In both Fig 4a and 4b the algorithm is only able to simulate data for the first two isotopes in the decay chains in 10 years, in each of those cases the radioactive parents and daughters undergo alpha decay which leads for confusion in the model when training as the program takes a sum for each of the types of decay in the decay chain. All other radioisotopes in Fig 1 are accurately identified.

Conclusion:

- The result from this example case shows that in some circumstances you can identify radioisotopes with machine learning. Future work will hopefully help the code evolve and make predictions of radioisotopes in the real world.

(GitHub repository)



(Digital Poster)