## **About Dataset**

The Fossil dataset was created to provide a comprehensive and realistic foundation for training and evaluating machine learning models aimed at predicting fossil ages. The dataset is of intermediate difficulty and includes a variety of geological, chemical, and physical attributes that are significant in the study of fossil formation and preservation.

### **Sources**

The initial data was sourced primarily from [PaleoBioDB](https://paleobiodb.org/" \l "/" \t "/Users/sarojrai/Desktop/my_doc/x/_blank), with additional private sources contributing to the dataset. After creating a small, initial dataset, a deep learning model was employed to expand and generate a synthetic version. This synthetic dataset simulates realistic scenarios, making it a valuable tool for data scientists and researchers in the field.

#### **Features**

****uranium\_lead\_ratio:**** Ratio of uranium to lead isotopes in the fossil sample.

****carbon\_14\_ratio:**** Ratio of carbon-14 isotopes present in the fossil sample.

****radioactive\_decay\_series:**** Measurement of the decay series from parent to daughter isotopes.

****stratigraphic\_layer\_depth:**** Depth of the fossil within the stratigraphic layer, in meters.

****isotopic\_composition:**** Proportion of different isotopes within the fossil sample.

****fossil\_size:**** Size of the fossil, in centimeters.

****fossil\_weight:**** Weight of the fossil, in grams.

****geological\_period:**** Geological period during which the fossil was formed.

****surrounding\_rock\_type:**** Type of rock surrounding the fossil.

****paleomagnetic\_data:**** Paleomagnetic orientation data of the fossil site.

****stratigraphic\_position:**** Position of the fossil within the stratigraphic column.

****age:**** Calculated age of the fossil based on various features, in years.