**The Toxicity Prediction Challenge II**

**Kaggle**:

<https://www.kaggle.com/code/ganeshvarma6669/x2022gvd-toxicitypredicitoncode>

**GitHub Link:**

<https://github.com/GaneshRuddarraju/Toxicity_Prediction_202206894>

**Google drive Link:**

<https://drive.google.com/drive/folders/10-nhCqogfqLQf2W6eILwzCr82Y9zZRpD?usp=share_link>

Install anaconda and packages:

* Download Anaconda
* Install Anaconda on your system
* Open Anaconda Navigator
* Open Jupyter Notebook from Anaconda Navigator
* Open Conda Terminal and run below commands

1. pip install rdkit
2. pip install CatBoost
3. pip install sklearn
4. Version of Anaconda Jupyter Server Information:

The version of the notebook server is: **6.4.8**  
The server is running on this version of Python: **3.9.12**

* Dataset and Code are available in the below link provided
* All the files are attached and also are also uploaded to drive link which is provided above.

**Steps to run the file:**

* **train\_II and test\_II** are the input csv files for the **Descriptors.ipynb** file.
* **desc\_test\_I.csv** and **desc\_train\_I.csv**  are generated as an output of the code **Descriptors.ipynb.**
* To run the descriptor file open **Descriptors.ipynb** file and click on cell and click on run all command.(approx. time. 3hrs)
* Now, run the **Toxicity\_draft.ipynb** file by keeping **desc\_test\_I.csv** and **desc\_train\_I.csv files** in the same path where we have all the python files available.
* Now, click on cell and click on run all command.

Graphical user interface, text, application

Description automatically generated