**Develop an AI prediction model for drug side effects using a common data model**

As part of the development of a drug side effect clue detection algorithm, the purpose of creating a time series prediction model that can evaluate the contribution of a specific variable at a specific point in time using a common data model.

side effects (drug) of interest

* Hepatotoxicity(valproic acid/lamotrigine) / nephrotoxicity(meloxicam/celecoxib)

Recommended H/W Specifications

* RAM 16GB or 32GB / GPU is recommended

S/W Requirements

* Environment where DBMS is installed.
* Permission to create a user table with SQL in the DB
* Install required packages python

Project file description

* \_log : Generate logs related to debug/errors that occur during execution
* \_sql : SQL statement to create DB table
* \_util : Python class functions used by internal code within scripts
* 0\_readDB : Create a table of cohorts and patients taking drugs in the DB.
* 1\_importSQL : Classify side effects by reading tables from DB
* 2\_preprocessing\_xgboost : Data preprocessing to run the XGBoost prediction model
* 3\_xgboost\_classification : Create XGBoost prediction model and evaluate performance metrics
* 4\_preprocessing\_lstm : Data preprocessing to run the BILSTM Attention model
* 5\_bi-lstm\_attention\_classification : Create Bi-lstm\_attention and evaluate performance
* 9\_code\_data\_visualization : Lab Test Data Distribution Visualization
* Data : directory for saving data that is created step by step
* Result : directory for saving performance metrics and figures

Installing

* Install project-related requirements in Python (If necessary, create a virtual environment)
* pip install -r requirements.
* graphviz install (https://graphviz.org/download/)
  + Check installation : cmd or terminal > "dot -V"
* pip install psycopg2-binary

Getting Started

* edit config.json file
  + 'working\_date' : Date to run the program.
  + 'dbms' : mssql or postgres
  + 'mssql' or 'postgresql' : server / user /password / port ..
  + 'meas' : meas\_concept\_id (used by the institution)
  + 'translatequerysql' : cdm\_database\_schema / target\_database\_schema / target\_database\_schema

Executing program

* run python script
* 2-1) Execute full script :
  + python main.py Step by step run with (y)/(n).
* 2-2) Run individual scripts :
  + cd 0\_cohort\_json python 0\_create\_cohort\_person\_in\_db.py