

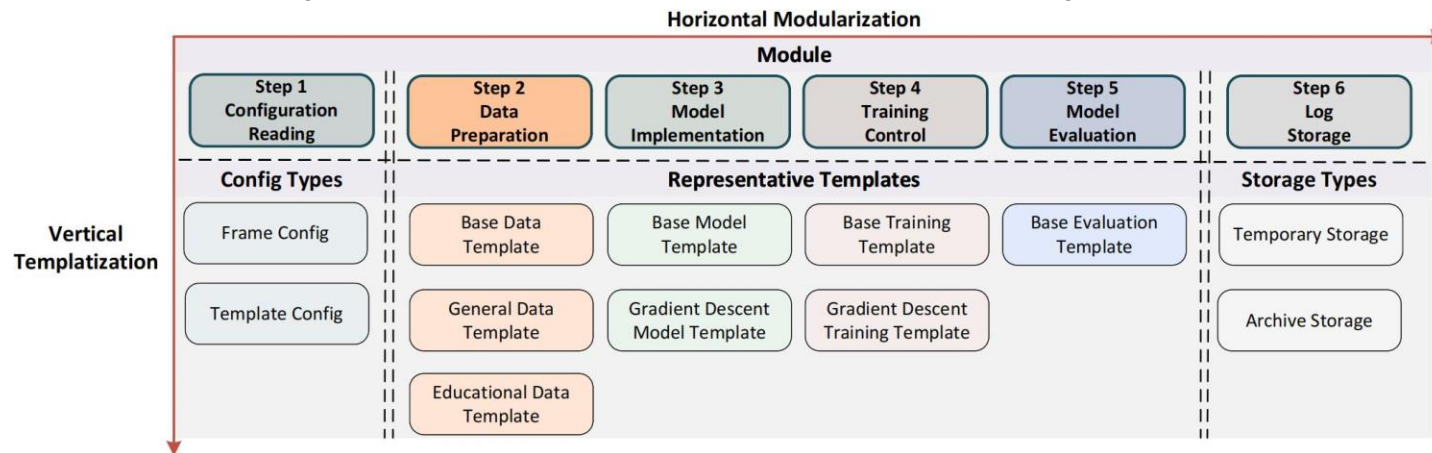
EduStudio: Towards a Unified Library for Student Cognitive Modeling

Le WU, Xiangzhi CHEN, Fei LIU, Junsong XIE, Chenao
XIA, Zhengtao TAN, Mi TIAN, Jinglong LI, Kun ZHANG,
Defu LIAN, Richang HONG, Meng WANG

Frontiers of Computer Science, DOI: [10.1007/s11704-024-40372-3](https://doi.org/10.1007/s11704-024-40372-3)

Problems & Ideas

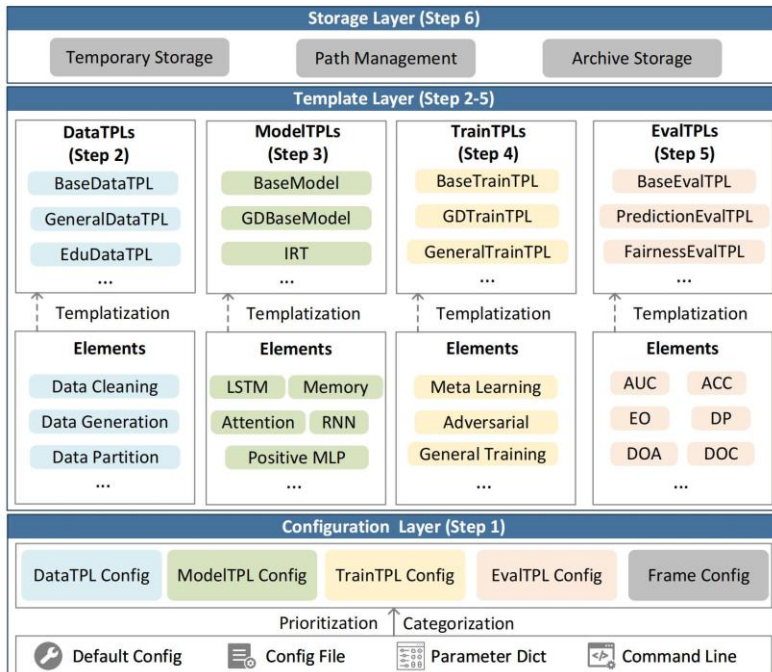
- Background and Problems:
 - Student cognitive modeling is a fundamental task in the intelligence education field, where Cognitive Diagnosis (CD) and Knowledge Tracing (KT) are two mainstream categories.
 - Existing open-source libraries often focus solely on specific categories, overlooking the intrinsic connections between CD and KT, and falling short in terms of reusability and flexibility.
- Ideas: We developed the **EduStudio** Library unifying CD and KT based on Pytorch. EduStudio employs the horizontal modularization and vertical templatization for better flexibility and reusability.



Design philosophy of EduStudio. Horizontal modularization: we decompose the general algorithmic pipeline into six modules to enhance flexibility. Vertical templatization: we implement reusable templates within the modules for steps 2-5 to achieve high-level management of complex elements.

Main Contributions

- Contributions:
 - The EduStudio integrates CD and KT models in a unified framework;
 - The EduStudio employs the horizontal modularization and vertical templatization, which provides better flexibility and reusability;
 - Relevant eco-services are provided surrounding EduStudio (eco-repository and leaderboard), which can enable more researchers to understand and quickly participate in student cognitive modeling.



The overall architecture of EduStudio

```
from edustudio.quickstart import run_edustudio
from edustudio.model import GDBaseModel
from edustudio.traintpl import GDBaseTrainTPL

class MyModel(GDBaseModel): # inherit basic model template
    default_cfg = {'emb_size': 64, 'lambda': 1.0}
    ...

class MyTrainTPL(GDBaseTrainTPL): # inherit basic training template
    default_cfg = {'batch_size': 64}
    ...

run_edustudio(
    dataset='assist-0910', # specify dataset name
    cfg_file_name=None, # specify configuration filename
    traintpl_cfg_dict={
        'cls': MyTrainTPL, # specify customized training template
        'batch_size': 128 # specify batch size of MyTrainTPL
    },
    datatpl_cfg_dict={
        'cls': 'CDInterDataTPExtendsQ' # specify name of data template
    },
    modeltpl_cfg_dict={
        'cls': MyModel, # specify customized model
        'emb_size': 32 # specify embedding size of MyModel
    },
    evaltpl_cfg_dict={
        # specify evaluation templates
        'cls': ['PredictionEvalTPL', 'InterpretabilityEvalTPL'],
        # specify used metrics in InterpretabilityEvalTPL
        'InterpretabilityEvalTPL': {'use_metrics': ['auc', 'acc']}
    },
    frame_cfg_dict={
        'CFG_FOLDER_PATH': '/home/edustudio/conf', # folder path of config file
    }
)
```

The code example of EduStudio