

Legal Commitment for Using the EduMix-24 and EduMath-24 Datasets

Licensor: Guangdong Institute of Smart Education, Jinan University,
Panyu District, Guangzhou, China

User: _____ **Institution:** _____
Address: _____

The Guangdong Institute of Smart Education created the EduMix-24 and EduMath-24 Datasets for Classifying Long-Content Educational Resources with Imbalanced Category Distributions. To ensure the correct and legal use of the EduMix-24 and EduMath-24 Datasets and prevent incidents such as data leakage, the User intends to enter into this Legal Commitment. The User is willing to be legally bound and promises to be subject to the following terms of use:

1. The Licensor grants the User the right to use the Dataset, but not the ownership. The Licensor retains the copyright of all sample data in the Dataset.
2. All samples in the Dataset in this commitment can only be used by the team of the named User and can only be used for academic research purposes. No samples can be used for any commercial purpose. In addition, the User shall not use the EduMix-24 and EduMath-24 Datasets for purposes that are not expressly set forth in this commitment, except with the written permission of the Licensor.
3. The User shall abide by the data protection regulations stipulated by the laws of the People's Republic of China, and is required to comply with all data protection laws that may affect him/her. The User shall not use the Dataset for any illegal activity (generating signature forgery, breaching users' identifiable information, etc.) and shall bear all legal responsibilities; otherwise, the User shall be liable for all legal consequences and the authorization will be permanently revoked.
4. The User is required to list 1-2 of the publications in recent 6 years to indicate that he/she (or the team) do research in the related research fields of Smart Education, Long-Content Classification, Imbalanced Categories, Prior Knowledge, Large Language Model, etc. in the application form.
5. The signature of the leader and the official seal of the institution (university, research institute, related department in the company, etc.) to which the User belongs must be provided;

otherwise, the application will be rejected.

6. The User must not copy or disseminate the Dataset without authorization and no part of the Dataset may be reproduced or distributed in any form or any means.
7. Considering that the Dataset contains personally identifiable information, the time series and images in the Dataset, and images rendered from time series, shall not be included in a publication unless expressly authorized in written by the Licensor.
8. All usages of the Dataset should be appropriately acknowledged. In publications, the Dataset should be referred to “EduMix-24 and EduMath-24 Datasets”. Publications of research results related to the Dataset should also be appropriately acknowledged.

Signature (User): _____

Seal (User): _____

Date: _____