

Re: movielens dataset request form

Daniel Kluver <kluve018@umn.edu>

Fr 11.11.2022 16:18

ACM ICDCN 2023

An:Eichinger, Tobias Sul Gi <tobias.eichinger@tu-berlin.de>;

Fantastic,

With that said, you have been approved to use the datasets as you have described in this form/email chain (educational or academic research, NON COMMERCIAL USE). Should any usage details change, you will need to fill out the form again. Please notify us of any publications that result from this usage, and follow the guidelines laid out for usage in the associated readme file.

As described in the README file, please acknowledge your use of MovieLens Datasets in any publication or demonstration by citing this paper: F. Maxwell Harper and Joseph A. Konstan. 2015. The MovieLens Datasets: History and Context. ACM Transactions on Interactive Intelligent Systems (TiiS) 5, 4: 19:1–19:19. <https://doi.org/10.1145/2827872> Have a great week, and happy researching! ~Kluver

On Fri, Nov 11, 2022 at 1:52 AM Eichinger, Tobias Sul Gi <tobias.eichinger@tu-berlin.de> wrote:

Dear Mr. Kluver,

thank you for your kind reply.

Your proposal works very well for me. Thank you!

Best wishes from Berlin,

Tobias

Tobias Eichinger, M.Sc.
Research assistant

Technische Universität Berlin | T-Labs
Chair of Service-centric Networking (SNET)

Ernst-Reuter-Platz 7, 10587 Berlin, Germany

Phone: +49 30 8353 58348
Mail: tobias.eichinger@tu-berlin.de
Web: <https://tu.berlin/snet/tobias-eichinger/>

ORCID: 0000-0002-8351-2823

Von: Daniel Kluver <kluve018@umn.edu>

Gesendet: Donnerstag, 10. November 2022 22:27:01

An: Eichinger, Tobias Sul Gi

Betreff: movielens dataset request form

Howdy,

I'm emailing in regards to your request to redistribute the movielens dataset.

This is not something we typically allow. There's a few reasons for this which all ultimately come down to being good stewards of our user's data -- if we for some reason needed to stop offering this data or make edits to comply with regulations we wouldn't want there to be multiple versions accessible online. We've seen this happen with other datasets in the past where different very slightly different versions of datasets exist due to redistribution similar to what you've outlined.

The normal "workaround" for cases like yours is to upload a script of some sort which will automatically download the dataset in question (we maintain stable URLs so this can be done even in archival services like Zenodo) and automatically extracts and processes it. In that way you still have an essentially complete and scripted process for replication, but you do not create a secondary point for people to obtain a copy of the MovieLens 25M dataset.

Would a solution like this work for you?

~Daniel

Kluver