

Appendix of Can We Trust Recommender System Fairness Evaluation? The Role of Fairness and Relevance

Theresia Veronika Rampisela

thra@di.ku.dk

University of Copenhagen
Copenhagen, Denmark

Maria Maistro

mm@di.ku.dk

University of Copenhagen
Copenhagen, Denmark

Tuukka Ruotsalo

tr@di.ku.dk

University of Copenhagen
Copenhagen, Denmark
LUT University
Lahti, Finland

Christina Lioma

c.lioma@di.ku.dk

University of Copenhagen
Copenhagen, Denmark

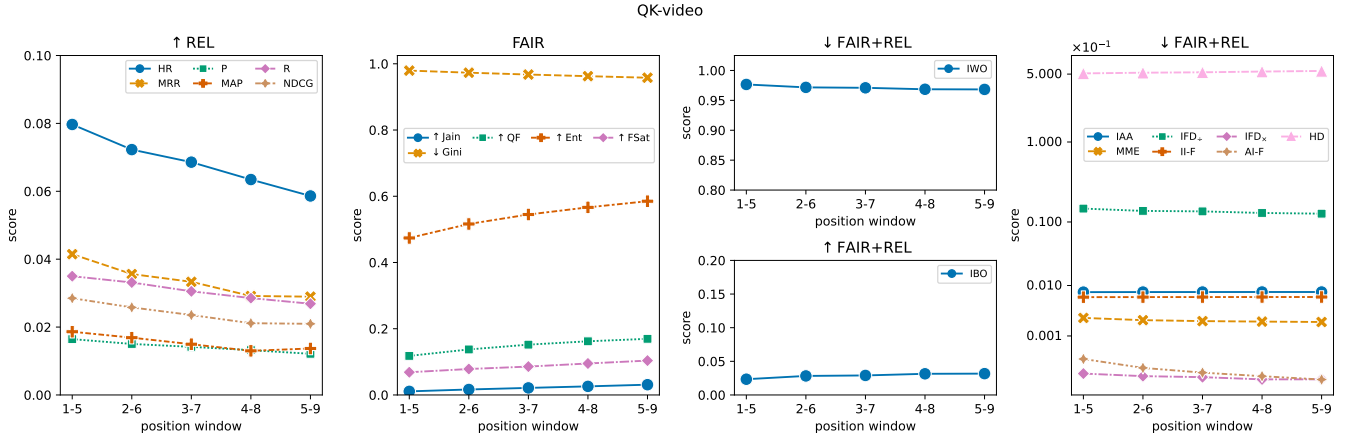


Figure 1: Sliding window evaluation ($k = 5$) of NCL for QK-video. The last column is in exponential scale.

ABSTRACT

This is the appendix of the paper “Can We Trust Recommender System Fairness Evaluation? The Role of Fairness and Relevance” [1].

CCS CONCEPTS

• **Information systems** → **Evaluation of retrieval results; Recommender systems**; • **General and reference** → **Evaluation**.

KEYWORDS

fairness and relevance evaluation; recommender systems

A EXTENDED RESULTS FOR MEASURE SENSITIVITY AT DIFFERENT RANKS

We present the results of our analysis on measure sensitivity at different ranks (RQ3) for QK-video in Fig. 1.

REFERENCES

- [1] Theresia Veronika Rampisela, Tuukka Ruotsalo, Maria Maistro, and Christina Lioma. 2024. Can We Trust Recommender System Fairness Evaluation? The Role of Fairness and Relevance. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. DOI: 10.1145/3626772.3657832.



This work is licensed under a Creative Commons Attribution-ShareAlike International 4.0 License.

SIGIR '24, July 14–18, 2024, Washington, DC, USA.

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0431-4/24/07

<https://doi.org/10.1145/3626772.3657832>