

Association for Information Systems

AIS Electronic Library (AISeL)

NEAIS 2022 Proceedings

New England Chapter of Association for
Information Systems

10-29-2022

An Empirical Study to Uncover Data Bias and Subsequent Algorithmic Bias in Recommendation

Esi Adeborna

University of Massachusetts Lowell, esi_adeborna@student.uml.edu

Follow this and additional works at: <https://aisel.aisnet.org/neais2022>

Recommended Citation

Adeborna, Esi, "An Empirical Study to Uncover Data Bias and Subsequent Algorithmic Bias in Recommendation" (2022). *NEAIS 2022 Proceedings*. 17.

<https://aisel.aisnet.org/neais2022/17>

This material is brought to you by the New England Chapter of Association for Information Systems at AIS Electronic Library (AISeL). It has been accepted for inclusion in NEAIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

An Empirical Study to Uncover Data Bias and Subsequent Algorithmic Bias in Recommendation

Completed Research Paper

Esi Adeborna
University of Massachusetts, Lowell
esi_adeborna@uml.edu

ABSTRACT

With the increase in the use of algorithms in decision making in different spheres of life, the need for “fairness” in these machine learning (ML) algorithms has become a progressively important concern. However, any remedy to bias must start with awareness that bias exists as it helps guide toward a solution. Therefore, this work is an empirical study that aims to uncover data and subsequent algorithmic bias in one of such ML applications. Specifically, we consider recommender systems (RS), one of the most successful applications of ML technology in practice today, and answer three main questions: (1) Does bias exist in RS input data? (2) Does the use of RS promote diversity of recommendation list or reinforces existing data bias? and (3) Does the application of RS over the years improve the recommendation diversity and reduce bias? We use Cell Phones and Accessories purchase dataset from Amazon for our empirical study. We employ the Long Tail phenomenon and quantify the shape of purchase distribution by calculating Gini coefficient from the Lorenz curve to determine the presence of bias. In the end, our findings from our experiments showed the presence of bias in both data and the RS algorithm year-over-year.

Keywords

Data Bias, Algorithmic Bias, Machine Learning, Gini Coefficient, Long tail, Recommender Systems.