Conference Abstract

How Reproducible are the Results Gained with the Help of Deep Learning Methods in Biodiversity Research?

Waqas Ahmed[‡,§], Vamsi Krishna Kommineni[‡,|,¶], Birgitta Koenig-ries[‡,#,|], Sheeba Samuel[#,‡]

‡ Friedrich Schiller University Jena, Jena, Germany
§ Thuringian Center for Learning Systems and Robotics, Ilmenau, Germany
| German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany
¶ Max Planck Institute for Biogeochemistry, Jena, Germany
# Michael Stifel Center Jena, Jena, Germany

Abstract

In recent years, deep learning methods in the biodiversity domain have gained significant attention due to their ability to handle the complexity of biological data and to make processing of large volumes of data feasible. However, these methods are not easy to interpret, so the opacity of new scientific research and discoveries makes them somewhat untrustworthy. Reproducibility is a fundamental aspect of scientific research, which enables validation and advancement of methods and results. If results obtained with the help of deep learning methods were reproducible, this would increase their trustworthiness. In this study, we investigate the state of reproducibility of deep learning methods in biodiversity research.

We propose a pipeline to investigate the reproducibility of deep learning methods in the biodiversity domain. In our preliminary work, we systematically mined the existing literature from Google Scholar to identify publications that employ deep-learning techniques for biodiversity research. By carefully curating a dataset of relevant publications, we extracted reproducibility-related variables for 61 publications using a manual approach, such as the

availability of datasets and code that serve as fundamental criteria for reproducibility assessment. Moreover, we extended our analysis to include advanced reproducibility variables, such as the specific deep learning methods, models, hyperparameters, etc., employed in the studies.

To facilitate the automatic extraction of information from publications, we plan to leverage the capabilities of large language models (LLMs). By using the latest natural language processing (NLP) techniques, we aim to identify and extract relevant information pertaining to the reproducibility of deep learning methods in the biodiversity domain. This study seeks to contribute to the establishment of robust and reliable research practices. The findings will not only aid in validating existing methods but also guide the development of future approaches, ultimately fostering transparency and trust in the application of deep learning techniques in biodiversity research.

Keywords

reproducibility, large language models (LLMs)

Presenting author

Waqas Ahmed

Presented at

TDWG 2023

Conflicts of interest

The authors have declared that no competing interests exist.