

# Investigating reproducibility of autoML-driven research papers



## ➤ Problem setting

The rise of automated machine learning (autoML) tools enables researchers to build their own machine learning (ML) models without extensive ML expertise. While these tools offer high levels of usability by hiding and automating complex and repetitive tasks, they also tend to lack transparency when it comes to their internal processes. While traditional ML tools and especially neural networks are already considered opaque black boxes the introduction of autoML increases this lack of transparency even further. This raises the question of whether these tools are suitable for conducting scientific research that demands high levels of documentation, transparency and reproducibility.

## ➤ Task description

- Identify scientific studies that use autoML tools as (part of) their research method
- Re-implement the same studies and try to reproduce the initial results
- Investigate why the reproduction was either possible or not

## ➤ Supervision

- Supervisor – Uni Graz: Univ.-Prof. Dr. Stefan Thalmann
- Supervisor – TU Graz (for CSS students): Anthony Ross-Hellauer, MA MSc Ph.D.
- Co-Supervisor : Armin Haberl, MSc BSc

**Business Analytics and**

**Data Science-Center**

**BANDAS-Center**