A strawman with machine learning for a brain: A response to Biedermann (2022) the strange persistence of (source) “identification” claims in forensic literature

A R T I C L E   I N F O

Keywords
Forensic inference
Machine learning

A B S T R A C T

We agree wholeheartedly with Biedermann (2022) FSI Synergy article 100222 in its criticism of research publications that treat forensic inference in source attribution as an “identification” or “individualization” task. We disagree, however, with its criticism of the use of machine learning for forensic inference. The argument it makes is a strawman argument. There is a growing body of literature on the calculation of well-calibrated likelihood ratios using machine-learning methods and relevant data, and on the validation under casework conditions of such machine-learning-based systems.

Letter to Editor:

Biedermann [1] is critical of research publications that treat forensic inference in source attribution as an “identification” or “individualization” task. Biedermann [1] argues that such publications condone unscientific attitudes and practices, foster unrealistic expectations among consumers of forensic science, and undermine trust in peer-reviewed publications because so-called “original research papers” are not, in fact, well grounded. With respect to these points, we agree wholeheartedly with Biedermann [1].

With respect to criticism of machine learning, however, we feel that Biedermann [1] makes a strawman argument. It defines “standard” machine learning as outputting categorical decisions and then criticizes the use of “standard” machine learning for forensic inference because it outputs categorical decisions. There are indeed research publications that misapply machine learning to forensic-inference problems, including using algorithms that output categorical decisions, e.g. [2]. But we fear that many readers will get the impression from Biedermann [1] that this is the only way (or at least the primary way) that machine learning is applied to forensic inference. There is in fact a growing body of literature on the calculation of well-calibrated likelihood ratios using machine-learning methods and relevant data, and on the validation under casework conditions of such machine-learning-based systems. Recent examples include [3–11].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Morrison, Ramos, Ypma: Writing - Original Draft, Writing - Review & Editing. All other authors: Writing - Review & Editing.

Acknowledgements

The writing of this response was supported by Research England’s Expanding Excellence in England Fund as part of funding for the Aston Institute for Forensic Linguistics 2019–2023.

References


Geoffrey Stewart Morrison*  
Forensic Data Science Laboratory, Aston University, Birmingham, UK  
Forensic Evaluation Ltd, Birmingham, UK

Daniel Ramos  
AUDIAS – Audio, Data Intelligence and Speech, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain

Rolf JF Ypma  
Netherlands Forensic Institute, The Hague, the Netherlands

Nabanita Basu  
Netherlands Forensic Institute, The Hague, the Netherlands

Kim de Bie  
Netherlands Forensic Institute, The Hague, the Netherlands

Ewald Enzinger  
Eduworks Corporation, Corvallis, OR, USA

Zeno Geradts  
Netherlands Forensic Institute, The Hague, the Netherlands

Didier Meuwly  
Netherlands Forensic Institute, The Hague, the Netherlands

David van der Vloed  
Netherlands Forensic Institute, The Hague, the Netherlands

Peter Vergeer  
Netherlands Forensic Institute, The Hague, the Netherlands

Philip Weber  
Netherlands Forensic Institute, The Hague, the Netherlands

* Corresponding author.  
E-mail address: geoff-morrison@forensic-evaluation.net (G.S. Morrison).