Technische Hochschule Brandenburg
University of Applied Sciences

IARIA

INFERRING POLITICAL ORIENTATION FROM CREDIT SCORE-RELEVANT VARIABLES

An Empirical Study on Profiling and Proxy Inference Through Sensitive Attributes

David Schnepp, *Brandenburg University of Applied Science, Germany.*

Olga Levina, *Brandenburg University of Applied Science, Germany.*

Contact mail: schnepp@th-brandenburg.de

DAVID SCHNEPP

Education:

- Research Stay, *Icahn School of Medicine at Mount Sinai, New York City, USA*. Research focuses on how hyperparameter optimization impacts the results of various CCC-tools
- Master of Science, *Brandenburg University of Applied Science, Brandenburg a. d. Havel, Germany*. Business Informatics.
- Bachelor of Arts, *Berlin School of Economics and Law, Berlin, Germany*. Business Administration.

Work history:

- Local area manager, *Lidl, Berlin, Germany*. Managed multiple Lidl stores, with responsibility for staff supervision, operational performance, communication, and compliance with company standards.

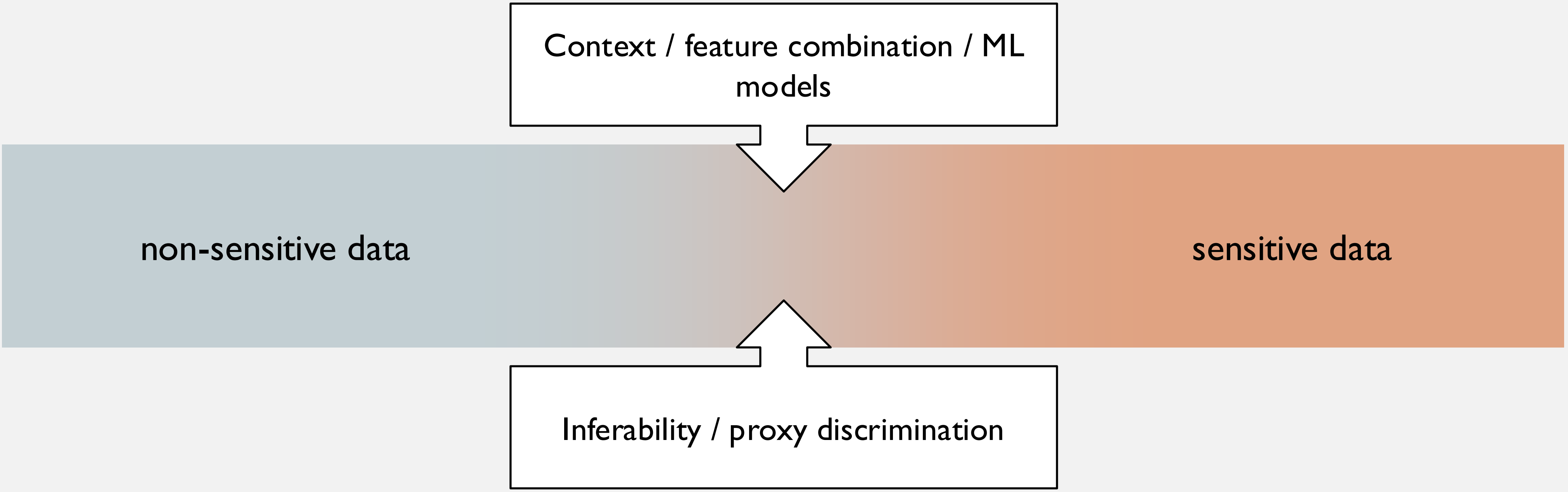
TOPIC & AIM

This paper…

- …discusses the risk that sensitive attributes may be inferred indirectly from seemingly non-sensitive data.
- …examines whether political orientation can be predicted from creditworthiness-related variables using supervised machine learning models in a specific setting.
- …assesses inferability of demographic and socioeconomic data in the context of the 2021 German federal election.
- …outlines implications for proxy-based profiling, data protection, and AI regulation.

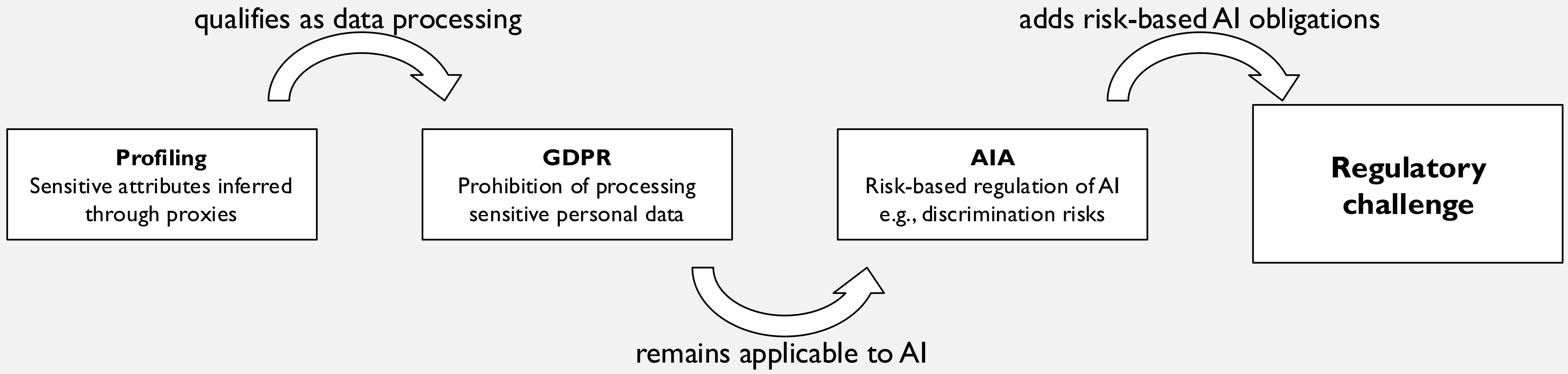
BACKGROUND

- Sensitive attributes do not need to be explicitly collected to become relevant for automated decision-making [1][3[4][5].
- This affects several domains such as finance, public administration, and political processes [1][2].
- The key concern shifts from the type of input data to the information that can be inferred from it [7][8].



BACKGROUND

- Political orientation is a protected sensitive attribute within the meaning of GDPR but may not need to be explicitly collected to become inferable [1][5][10].
- This creates a regulatory challenge where sensitive attributes are inferred rather than directly processed.
- Creditworthiness-related variables often include demographic and socioeconomic information that could potentially act as proxies.



METHODICAL APPROACH

Input [11]

Credit-Scoring related feature [16]	**Data Type** [16]
Gender	Nominal
Age	Metric
Domicile / housing location	Ordinal
Relationship status	Nominal
Household size	Metric
Children living in household	Metric
Contract type	Nominal
General education level	Ordinal
Vocational education level	Ordinal
Household income	Ordinal

European Social Survey Vol .10

machine learning models

Logistic Regression
Support Vector Machine
K-Nearest Neighbors
Boosting-based Decision Tree

Target [11]

Who did you vote for?
Federal election September 2021, Germany [16]

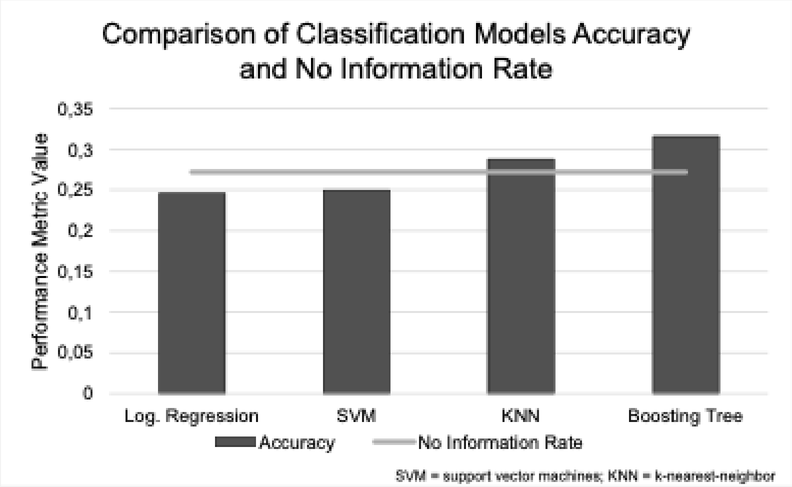
- ☐ CDU/CSU
- ☐ SPD
- ☐ Die Linke
- ☐ Bündnis 90/Die Grünen
- ☐ FDP
- ☐ AfD

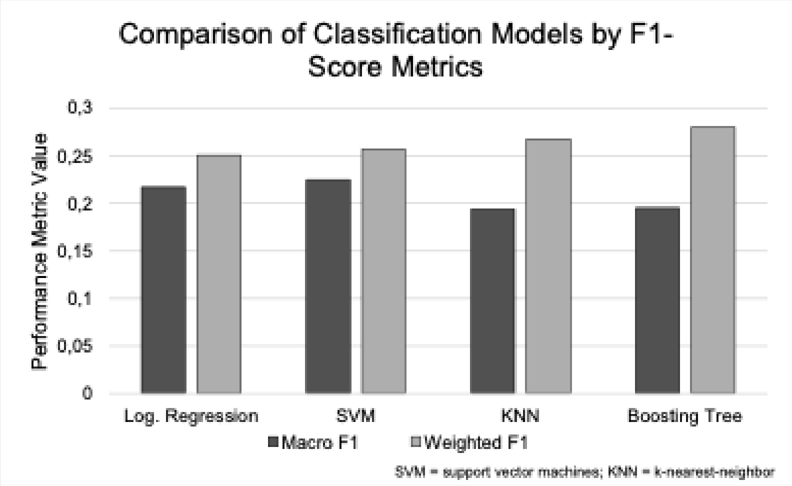
→ Multi-classification problem

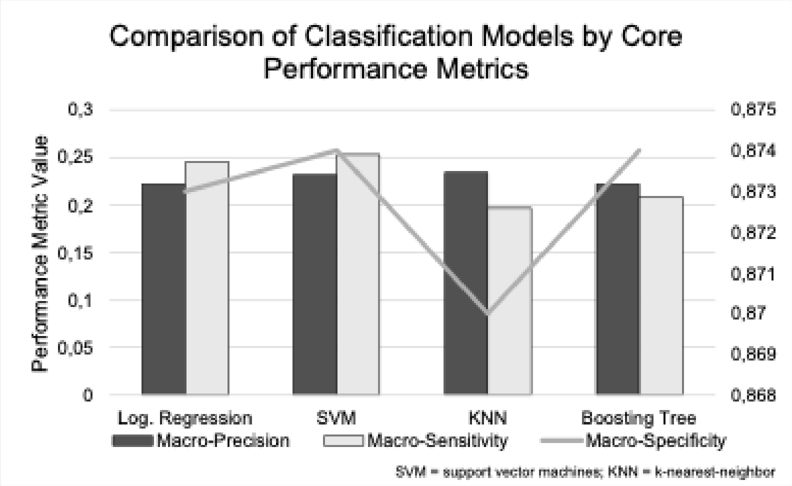
European Social Survey Vol .10

RESULTS

- No reliable prediction of political orientation is possible across all applied model types.
- A permutation-based feature-importance analysis showed that none of the selected creditworthiness-related variables had strong predictive relevance.







DISCUSSION

- **No meaningful inference was observed in this empirical setup.**
  Political orientation could not be reliably predicted from the selected creditworthiness-related variables.
- **The finding is context-specific and not generally transferable.**
  It does not prove that political orientation cannot be inferred from creditworthiness-related data in other datasets, populations, or model settings [12][13][14][15].
- **The reason for the negative result remains open.**
  Poor performance may reflect genuine non-inferability, but also methodological factors such as class imbalance, preprocessing choices, limited feature richness, or model selection.
- **This uncertainty creates a regulatory challenge.**
  Inferability cannot be reliably ruled out ex ante, making proxy-based profiling difficult to assess under GDPR and AI regulation.

OUTLOOK

1) Future research should move from binary risk assessments toward context-sensitive inferability evaluations that systematically examine how data richness, class distribution, model choice, and resampling strategies affect the feasibility of attribute profiling [12][13][14][15].

2) Since inferability appears to be attribute-specific, future work should examine which categories of sensitive information are particularly vulnerable to indirect inference and support this through benchmarks and evaluation protocols.

3) As absence of predictive performance is not evidence of absence, system design and policy should strengthen auditing methods, privacy-enhancing technologies, and data-minimization practices.

REFERENCES

[1] X. Gao, *C. Shen, W. Jiang, C. Lin, Q. Li, Q. Wang, Q. Li, and X. Guan,*, "Fairness in machine learning: definition, testing, debugging, and application", *Sci. China Inf. Sci.*, vol. 67, no. 9, p. 191201, Sep. 2024, doi: 10.1007/s11432-023-4060-x.
[2] Risk Research, "Der EU AI Act in Kreditinstituten", Aug. 2024. Accessed: May 09, 2026. [Online]. Available: https://www.risk-research.de/fileadmin/userdaten/docs/PDF-Dateien/Whitepaper/Risk_Research_Whitepaper_7_EU_AI_Act.pdf.
[3] P. Quinn and G. Malgieri, "The difficulty of defining sensitive data—the concept of sensitive data in the EU data protection framework", *Ger. Law J.*, vol. 22, no. 8, pp. 1583–1612, Dec. 2021, doi: 10.1017/glj.2021.79.
[4] F. Elsafoury, "Diskriminierung | Wenn der Schein trügt – Deepfakes und die politische Realität", bpb.de. Accessed: May 09, 2026. [Online]. Available: https://www.bpb.de/lernen/bewegtbild-und-politische-bildung/556762/diskriminierung/.
[5] ARTIKEL-29-DATENSCHUTZGRUPPE, "Leitlinien zu automatisierten Entscheidungen im Einzelfall einschließlich Profiling für die Zwecke der Verordnung 2016/679", Oct. 2017. Accessed: May 09, 2026. [Online]. Available: https://dsb.gv.at/sites/site0344/media/downloads/wp251rev01_de.pdf.
[6] S. Yeom, A. Datta, and M. Fredrikson, "Hunting for discriminatory proxies in linear regression models", 2018, *arXiv*. doi: 10.48550/ARXIV.1810.07155.
[7] A. D. Selbst, D. Boyd, S. A. Friedler, S. Venkatasubramanian, and J. Vertesi, "Fairness and abstraction in sociotechnical systems", in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, Atlanta GA USA: ACM, Jan. 2019, pp. 59–68, doi: 10.1145/3287560.3287598.
[8] K. Tiwari, N. Sarkar, K. Bisht, and S. Shukla, "A systematic survey on bias and fairness in machine learning", in *Data Science and Security*, S. Shukla, H. Sayama, K. Tiwari, and J. V. Kureethara, Eds., Lecture Notes in Networks and Systems, vol. 1355. Singapore: Springer Nature Singapore, 2025, pp. 77–91, doi: 10.1007/978-981-96-4883-2_8.

REFERENCES

[9] European Data Protection Supervisor, *AI Act Regulation (EU) 2024/1689*. Luxembourg: Publications Office of the European Union, 2025, doi: 10.2804/4225375.

[10] European Parliament and Council of the European Union, *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*, May 4, 2016. Accessed: May 09, 2026. [Online]. Available: http://data.europa.eu/eli/reg/2016/679/oj.

[11] European Social Survey European Research Infrastructure (ESS ERIC), "ESS10 - data from Interviewer's questionnaire, edition 3.1". Sikt - Norwegian Agency for Shared Services in Education and Research, 2023, doi: 10.21338/ESS10INTE03_1.

[12] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior", *Proc. Natl. Acad. Sci.*, vol. 110, no. 15, pp. 5802–5805, Apr. 2013, doi: 10.1073/pnas.1218772110.

[13] B. K. Hassani, "Societal bias reinforcement through machine learning: a credit scoring perspective", *AI Ethics*, vol. 1, no. 3, pp. 239–247, Aug. 2021, doi: 10.1007/s43681-020-00026-z.

[14] R. Chaturvedi and S. Chaturvedi, "It's all in the name: a character-based approach to infer religion", *Polit. Anal.*, vol. 32, no. 1, pp. 34–49, Jan. 2024, doi: 10.1017/pan.2023.6.

[15] S. D. Kim, G. Andreeva, and M. Rovatsos, "The double-edged sword of big data and information technology for the disadvantaged: a cautionary tale from open banking", 2023, *arXiv*, doi: 10.48550/ARXIV.2307.13408.

[16] European Social Survey ERIC (ESS ERIC), "European Social Survey (ESS), Round 10 - 2020", 2022, *Sikt - Norwegian Agency for Shared Services in Education and Research*, doi: 10.21338/NSD-ESS10-2020.

[17] S. Selle, *Data Science Training - Supervised Learning: Ein praktischer Einstieg ins überwachte maschinelle Lernen*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2024, doi: 10.1007/978-3-662-67960-9.