Vollmer Group Expression of Interest

Here is the list of projects currently available in our group. If there is a project not listed here that you feel particularly passionate about, you can suggest your own topic and we will see if there is a researcher in the group that could support you. You can select up to 3 of our topics, and additionally suggest one of your own. This will help us match you with the right person in our group.

Starting from The Landscape and Gaps in Open Source Fairness Toolkits and https://dl.acm.org/doi/pdf/10.1145/3610536 and finding newer references survey the current approaches in performing well on multiple metrics with a focus on fairness.

2. Prior elicitation with LLMs

3. Extending multi-omics integration methods to different numerical scales

In biomedical research, a number of methods are available for continuous or binary data types, but transcripts are typically recorded as counts, disease-free survival of patients is encoded as time-to-event data and patient-reported outcomes or cancer stages are ordered categories. Can we extend machine learning-based integration methods to handle such modes? Reference https://journals.sagepub.com/doi/full/10.1177/1177932219899051 (see challenges)

4. Transformer models for human biobank data

Generative pre-trained transformers (GPT) are popular models for computer vision and natural language, applied to a whole host of practical applications. However, what if we want to train and/or apply such models to tabular datasets, such as electronic health records? To what extent can GPTs synthesize or impute realistic-looking patient records? What if the number of records available is relatively small? Reference [2306.11547] Event Stream GPT: A Data Pre-processing and Modeling Library for Generative, Pre-trained Transformers over Continuous-time Sequences of Complex Events

5. Biologically-informed feature selection for survival analysis from high-dimensional data

What effect does incorporating biological knowledge - such as pathway information - have on efforts to select features from high-dimensional datasets for survival tasks? This project compares classical regularization-based approaches with deep learning architectures such as pathway activity variational autoencoders. It may also investigate the extent to which semi-supervised learning may have on the ability of such an AE to find useful latent features. Reference Large-scale benchmark study of survival prediction methods using multi-omics data | Briefings in Bioinformatics | Oxford Academic Incorporating Prior Knowledge in Deep Learning Models via Pathway Activity Autoencoders Supervised autoencoders: improving generalization performance with unsupervised regularizers

6. Pathway activity autoencoder architectures

Multi-omics datasets are very high-dimensional, necessitating different methods of feature selection to reduce the dimensionality of the data and enable the fitting of predictive models and to make them simple enough to interpret in a biological context. Incorporating prior information can make the extracted or selected features more interpretable and requires fewer parameters or features than uninformed models. In this project you will review different ways in which pathway information is incorporated into unsupervised and supervised machine learning architectures for high-dimensional -omics data. Example references [2306.05813] Incorporating Prior Knowledge in Deep Learning Models via Pathway Activity Autoencoders Biologically informed deep neural network for prostate cancer discovery | Nature A fair experimental comparison of neural network architectures for latent representations of multi-omics for drug response prediction

7. Cross-species transfer learning on non-imaging data for cardiovascular disease

Transfer learning — i.e. pretraining a model on a large dataset for one task and then fine-tuning it on a separate task for which less data are available — is widely applied in machine vision and in natural language processing. Meanwhile in biomedical research, patient data are sensitive and expensive to collect, especially in multi-omics. To what extent can models trained on one context (e.g. another species) outperform models trained from scratch on human data? References Transfer learning efficiently maps bone marrow cell types from mouse to human using single-cell RNA sequencing | Communications Biology Transfer learning for non-image data in clinical research: A scoping review | PLOS Digital Health

8. Transfer learning on open data for cardiovascular disease

Transfer learning is widely applied in machine vision and in natural language processing. Meanwhile in biomedical research, data are sensitive and expensive to collect, especially in multi-omics, but a number of open benchmark datasets are available, such as Cancer Genome Atlas project (TCGA) or MIMIC, as well as large biobanks like the UK Biobank. However, most research in this area focuses on oncology rather than cardiovascular disease (CVD). Given the popularity of unsupervised or self-supervised learning (e.g. VAEs) for learning representations of high-dimensional biomedical data, it may follow that embeddings learnt from high-data settings may be useful in low-data settings, even in a different context. Thus the aim of this project is to investigate the extent to which transfer learning can aid research into multi-omic CVD research, both in principle and in practice, using a combination of openly available CVD and non-CVD datasets. References A Review of Integrative Imputation for Multi-Omics Datasets Mapping single-cell data to reference atlases by transfer learning | Nature Biotechnology

9. Knowledge graph neural networks for multi-modal biological entities in cardiovascular disease

Domain knowledge about biological entities, such as protein-protein interactions, drug interactions or genetic pathways, is often encoded in the form of knowledge graphs. Incorporating these data, especially those drawn from different sources, into a machine learning model, may not be straightforward. Previously, drug, gene and protein databases have been combined into knowledge graphs in the context of autoimmune diseases and cancer, however this is not an area yet widely explored in cardiovascular disease. This project explores the different ways in which pathway information can be incorporated into multi-omics models in the form of graph neural network architectures, for supervised and unsupervised learning problems. References Precision Medicine Oriented Knowledge Graph - Zitnik Lab The Impact of Pathway Database Choice on Statistical Enrichment Analysis and Predictive Modeling

 \wedge

 \wedge

 \wedge

 \wedge

 \wedge

 \wedge

 \wedge

 \wedge

 \wedge

 \checkmark

 \wedge

 \wedge

Discrete survival analysis, also known as 'survival stacking' or 'reduction' is a way of preprocessing time-to-event data that allows it to be cast as a classification problem. This allows generic machine learning frameworks to be used instead of limiting oneself to survival-specific models, which may not be able to scale to large or highdimensional datasets. Though this approach is not new, it is not well-described in the literature and tools for this purpose are hard to find or poorly documented. References Survival prediction models: an introduction to discrete-time modeling | BMC Medical Research Methodology [1909.11171] Survival analysis as a classification problem Deep Neural Networks for Survival Analysis Using Pseudo Values | IEEE Journals & Magazine Article A Fast and Accurate Method for Genome-Wide Time-to-Event Data Analysis and Its Application to UK Biobank

11. Prediction of Survey outcome in CMIE

In this project, we consider the task of 'extreme multi-label classification' applied to predicting responses to surveys and questionnaires. What methods can be used to predict missing/future survey responses (based on previous, future, nearby or similar surveys) and what metrics (e.g. macro F1 score) would help evaluate the quality of such predictions? What makes some parts of surveys harder to predict than others? The Consumer Pyramids Household Survey is a continuous survey of over 236,000 Indian households, collected by the Centre for Monitoring Indian Economy (CMIE). In this project you will explore the feasibility of applying different methods, as described above, to predicting and imputing such data and its implications on active survey.

12. Causal and explainable models for multiple survival prediction

Survival analysis is commonly used in medicine, epidemiology and industrial applications to model the temporal distribution of events. How can we integrate model explanations that might vary with time? In this project we combine ideas of explainable AI, online learning, discrete survival analysis and temporal cross-validation to examine dynamic survival predictions and their explanations. References Alicia Curth, Changhee Lee, Mihaela van der Schaar. SurvITE: Learning Heterogeneous Treatment Effects from Time-to-Event Data. NeurIPS, 2021. https://openreview.net/forum?id=f0_tkoEJV88 https://research.atspotify.com/publications/temporally-consistent-survival-analysis/ Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek. SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems, 2023 https://doi.org/10.1016/j.knosys.2022.110234

13. Visualization makeovers with LLMs

Data visualizations are an essential component to effectively communicating insights derived from quantitative information. However, making a 'good' graphic that accurately and efficiently conveys the desired message to the reader is both an art and a science, typically not taught in the data science curriculum. 'Visualization makeovers' are exercises where the (online) community of data analysts exchange feedback on improving graphical representations of a given dataset. Can LLMs emulate this task? Given a plot in the form of an image file, or the code used to generate it, can we use AI, via LLMs, to partially automatically generate constructive criticism to help produce a 'better' version? The primary method would be centered around prompt engineering and leveraging LLMs' latent knowledge on data visualizations. Instead of explicit training on labeled paired comparisons of graph types, it relies on the assumption that information about good and bad visualization practices is already stored within popular LLMs, e.g., ChatGPT, as discussed by people online. For evaluation, the challenge lies in establishing quantitative quality criteria for the generated critiques and descriptions. The project could consider metrics such as the clarity of descriptions and the extent to which the critique enhances the original visualization. Example references #MakeoverMonday LIDA: A Tool for Automatic Generation of Grammar-Agnostic Visualizations and Infographics using Large Language Models. Victor Dibia, 2023. https://doi.org/10.48550/arXiv.2303.02927 The Visual Display of Quantitative Information. Edward Tufte, 2001. https://www.edwardtufte.com/tufte/books_vdqi Using Large Language Models to Generate Engaging Captions for Data Visualizations. Ashley Liew and Klaus Mueller, 2022. https://doi.org/10.48550/arXiv.2212.14047

14. Analyzing longitudinal ordinal data as metric: what could possibly go wrong?

In biomedical research, mobile health (mHealth) describes the use of connected devices such as wearable sensors, smartwatches and smartphones, to collect health data during everyday experiences outside the clinical setting. Patients living with chronic diseases are often asked to report subjective symptoms such as pain, fatigue, mood or sleep quality via longitudinal questionnaires. These data are collected on an ordinal scale, e.g. choosing from labels such as "no pain", "moderate pain", "severe pain" and "worst pain imaginable". Treating cross-sectional ordered categorical data as continuous (i.e. replacing the above labels with integers 1-5) is a common domain practice that has the potential to result in erroneous estimates of treatment effects. To what extent is this also problem for longitudinal data, in the form of ordinal time series? Do machine learning models help or hinder the situation? References Analyzing ordinal data with metric models: What could possibly go wrong?

https://www.sciencedirect.com/science/article/abs/pii/S0022103117307746?via=ihub Extension of the following paper to time series / longitudinal data. https://www.sciencedirect.com/science/article/abs/pii/S0022103117307746?via=ihub Ordinal responses are one thing, but if we want an autoregressive model then there must also be a way of treating predictors as ordinal. One way of doing it is discussed in this paper. https://www.frontiersin.org/articles/10.3389/fams.2017.00015/full See also the papers that cite that one: https://scholar.google.com/scholar?cites=10582074246415117272&as_sdt=2005&sciodt=0,5&hl=en How can we model an ordinal time series that has (multiple) other ordinal time series (including lagged versions of itself) as predictors? Can we add in non-ordinal data as well, e.g. continuous longitudinal data, or baseline predictors? So we can maybe divide this up into sub-projects: Treating ordinal predictors as metric: what could possibly go wrong? Ordinal responses with ordinal features (static case). Ordinal time series as metric: what could possibly go wrong? The machine learning case (i.e. ordinal feature extraction versus ordinary feature extraction, for SVMs, random forests, NNs etc). Mixed random effects models? Individual trajectories? Relevant R packages include npreg and bigsplines, which allow constructing ordinal smoothing spline bases.

15. Active learning for survey design

16. Forensic archaeology for scientific reproducibility

If we are sceptical of a result published in the scientific literature, or curious about the methods used to reach it. Given the final result and access to the original data, but incomplete information on the methods used, can we use reinforcement learning and a library of reasonable data analysis steps to reconstruct a set of 'plausible' candidate data science pipelines that give the same (or similar) result? Example reference Query reverse engineering https://link.springer.com/article/10.1007/s00778-013-0349-3

17. Recommending journals for manuscripts

Given a draft manuscript, which academic journal would be the most appropriate place to publish it? That is, can we build a model, trained on open bibliometric data, that takes an article title and/or abstract as input, and generates a journal (or conference proceedings) recommendation as output? References https://thesai.org/Downloads/Volume8No12/Paper_27-Recommender_System_for_Journal_Articles.pdf https://www.emerald.com/insight/content/doi/10.1108/DTA-08-2019-0135/full/html

 \wedge

18. Contextual Bayesian non-stationary bandits for treatment recommendation

Motivated by the applications of clinical trials, where machine learning models can suggest treatments to clinicians and help patients ease symptoms more quickly. We start with IntelligentPooling [1], and study a contextual Bayesian non-stationary bandit model for personalised recommendation, where we capture the prior knowledge given by experts (e.g. Fig 2 in [9]) and contextual information about patients [96]. One feature in demand is to also suggest the treatment length given the non-stationary environment. We will study methods to handle this by considering the combination of treatment types and lengths as arms. The goal is to maximise the cumulative patient recovery rate by suggesting treatment type and lengths in each round. The performance will be evaluated empirically on simulated data and theoretically of regret bounds. Start Points: Start with the contextual bandits library https://github.com/david-cortes/contextualbandits and get familiar with contextual bandits algorithms Reproduce IntelligentPooling [1] and identify research gaps (strengths and weakness). Consider the combination of treatment types and lengths as arms, design novel algorithms Evaluation of proposed algorithm: 1) regret bounds 2) simulations, compare with intelligentPooling 3) real-world experiments [1] S. Tomkins, P. Liao, P. Klasnja, and S. Murphy, "IntelligentPooling: Practical Thompson Sampling for mHealth," Jul. 2020, doi: 10.48550/arXiv.2008.01571. [9] T. Gärtner, J. Schneider, B. Arnrich, and S. Konigorski, "Comparison of Bayesian networks, G-estimation and linear models to estimate causal treatment effects in aggregated N-of-1 trials," bioRxiv, Jul. 21, 2022. doi: Q. Wu, N. Iyer, and H. Wang, "Learning Contextual Bandits in a Non-stationary Environment," in The 41st International ACM SIGIR 10.1101/2022.07.21.22277832. [96] Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA, Jun. 2018, pp. 495–504.

19. Discovery and Synthesis of Data Science Pipelines via Agent Self-Experimentation

Existing tools that assist end-users in the design of data science pipelines are usually restricted to the scope of pipelines they are capable of suggesting. Depending on the implementation of such assistant, it can be restricted to a set of known pipelines (e.g., case-based intelligent discovery assistants), a set of rules and reasoning logic on how new pipelines are composed (e.g., HTN planners), or a meta-template of the data science pipeline that is initialized with a specific set of hyperparameters (e.g., most AutoML solutions). The common drawback of these tools: they cannot go "off script" and suggest anything that goes beyond their current state of knowledge representation. In this project, we explore approaches for discovery and synthesis of data science pipelines that allow experimentation (incl. making mistakes and creating pipelines that yield software exceptions) and enable "going off script", potentially extending the state of its knowledge over time. The example approaches are analogous to the field of automated feature engineering. where the base operators-"building blocks" are combined with the rules of operator composition and can generate potentially-unbound, complex feature processing functions; knowledge-intensive approaches that combine pipeline generation rules with data science ontologies; wisdom of the crowd approaches that accumulate a large set of known data science pipelines from multiple end-users to depict an ever-growing set of known composition rules in a data-driven way. Note: depending on the selected family of approaches, this project can be executed by multiple students (separately).

20. Development of an open-source software library for simulating patient data for N-of-1 trials

The purpose of this MSc thesis project is to develop an open-source software library for simulating patient data for N-of-1 trials. The library will capture important aspects of reality for the development and benchmarking of new algorithms. The existing work contributes to capture effects such as baseline drift, dropout, and carryover, but it does not capture all the complexities of our use cases. Therefore, this project will collaborate with clinicians to incrementally add features to develop realistic patient simulations. The library will include basic features such as different sequences and lengths of treatments, carry-over effects, drop-out effects, and missing data. Additionally, we waill aim include novel features such as unobserved confounders, limited intervenable nodes, nonstationary models, non-compliance with suggested treatment, and models with a high number of variables as a test bed for neuro-symbolic algorithms for explainable AI. These features will be developed separately and in a modular fashion, allowing experimenters to create custom test beds for their experiments. The work package is open source, allowing other contributors to add missing features and create a library of the most important features for N-of-1 trials. Alongside building the functionality, the project will create multiple benchmarks for the literature. The project's ultimate goal is to build the foundation provide researchers with a comprehensive and reliable tool for simulating patient data that can aid in the development and evaluation of new algorithms. Simulation of patients symptoms and effect of treatments that includes different sequences and lengths of treatments, carry-over effects, drop-out effects, and missing data. References T. Gärtner, J. Schneider, B. Arnrich, and S. Konigorski, "Comparison of Bayesian networks, G-estimation and linear models to estimate causal treatment effects in aggregated N-of-1 trials," bioRxiv, Jul. 21, 2022. doi: 10.1101/2022.07.21.22277832. [91] D. Kaur et al., "Application of Bayesian networks to generate synthetic health data," J. Am. Med. Inform. Assoc., vol. 28, no. 4, pp. 801–811, Mar. 2021. [92] A. Tucker, Z. Wang, Y. Rotalinti, and P. Myles, "Generating high-fidelity synthetic patient data for assessing machine learning healthcare software," NPJ Digit Med, vol. 3, no. 1, p. 147, Nov. 2020. [93] R. Horton, M. T. Hosseinabadi, and J.-M. Agosta, "Approaches to Optimizing Medical Treatment Policy using Temporal Causal Model-Based Simulation," Nov. 04, 2022. Accessed: Mar. 13, 2023. [Online]. Available: https://openreview.net/pdf?id=TptoTbkwaa [94] B. Percha, E. B. Baskerville, M. Johnson, J. T. Dudley, and N. Zimmerman, "Designing Robust N-of-1 Studies for Precision Medicine: Simulation Study and Design Recommendations," J. Med. Internet Res., vol. 21, no. 4, p. e12641, Apr. 2019.

21. Recovering individual trajectories from aggregated data: applications in health & society

Data aggregation is a common practice to protect privacy, but it often comes at the cost of losing information at the individual level. This work aims to develop methods that can recover joint distributions or even individual trajectories from only aggregated counts/rates or sparse data, leveraging weakly informative data such as micro-census or satellite data. We will extend the work of Singh et al. [94] on reconstructing event sequences based on count observations to continuous-time Markov chains. As a stretch goal, we will apply our methods to two applications: (1) breaking down the knowledge about people affected by multiple crises in Lesotho by multiple demographic attributes from aggregate separately collected data, and (2) breaking down internal migration by different groups using the global internal displacement database [95]. Our work will tie together both individual health and societal well-being and contribute to advancing the field of data aggregation. Objectives To review and compare existing methods for recovering individual trajectories from aggregated data, with a focus on continuous-time Markov chains and weakly informative data. To extend the work of Singh et al. [94] on reconstructing event sequences based on count observations to handle multiple demographic attributes and apply it to real-world datasets. To evaluate the performance of the proposed methods using simulated data and benchmark datasets, and compare them with existing methods. To apply the proposed methods to the two stretch goals of breaking down knowledge about people affected by multiple crises in Lesotho and internal migration by different groups. To discuss the implications and limitations of the proposed methods and suggest future research directions. Methodology Conduct a systematic literature review of existing methods for recovering individual trajectories from aggregated data, with a focus on continuous-time Markov chains and weakly informative data. Develop a framework for extending the work of Singh et al. [94] to handle multiple demographic attributes and apply it to real-world datasets. Evaluate the proposed methods using simulated data and benchmark datasets, and compare them with existing methods based on accuracy, efficiency, and scalability. Apply the proposed methods to the two stretch goals of breaking down knowledge about people affected by multiple crises in Lesotho and internal migration by different groups, and interpret the results in the context of health and society. Discuss the implications and limitations of the proposed methods, suggest future research directions, and provide recommendations for policymakers and practitioners. Expected outcomes A comprehensive review of existing methods for recovering individual trajectories from aggregated data, with a focus on continuous-time Markov chains and weakly informative data. A framework for extending the work of Singh et al. [94] to handle multiple demographic attributes and apply it to real-world datasets. A set of proposed methods for recovering individual trajectories from aggregated data that outperform existing methods based on accuracy, efficiency, and scalability. Applications of the proposed methods to breaking down knowledge about people affected by multiple crises in Lesotho and internal migration by different groups, providing new insights into health and society. Recommendations for policymakers and practitioners on how to use the References [92] Philip J. Brown and Clive D. Payne. Aggregate Data, Ecological Regression, and Voting Transitions. Journal of the American Statistical Association, 81(394):452–460, June 1986. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459. 1986.10478290. URL http://www.tandfonline.com/doi/abs/10.1080/01621459.1986.10478290. [93] Seth R. Flaxman, Yu-Xiang Wang, and Alexander J. Smola. Who Supported Obama in 2012?: Ecological Inference through Distribution Regression. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 289–298, Sydney NSW Australia, August 2015. ACM. ISBN 978-1-4503-3664-2. doi: 10.1145/2783258.2783300. URL https://dl.acm.org/doi/10.1145/2783258.2783300. [94] Rahul Singh, Qinsheng Zhang, and Yongxin Chen. Learning Hidden Markov Models from Aggregate Observations, November 2021. URL http://arxiv.org/abs/2011.11236. arXiv:2011.11236 [cs, eess]. [95] Tomoharu Iwata and Hitoshi Shimizu. Neural Collective Graphical Models for Estimating Spatio-Temporal Population Flow from Aggregated Data. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01):3935–3942, July 2019. ISSN 2374-3468, 2159-5399. doi: 10.1609/aaai.v33i01.33013935. URL https://ojs.aaai.org/index.php/AAAI/ article/view/4283. [96] Esther Rolf, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Vaishaal Shankar, Miyabi Ishihara, Benjamin Recht, and Solomon Hsiang. A generalizable and accessible approach to machine learning with global satellite imagery. Nature Communications, 12(1):4392, July 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-24638-z. URL https://www.nature.com/articles/s41467-021-24638-z.

 \wedge

 \wedge

 \wedge

 \checkmark

 \wedge

 \wedge

22. Preprocessing strategies for spatio-temporal event data

The aim of this MSc thesis project is to explore existing practices for cleaning, selecting, and extracting features from spatio-temporal event data, and to evaluate the sensitivity of temporal models to choices of data preprocessing strategies. The project will begin with a systematic review of commonly-used data cleaning methods and tools available, with a focus on their support for irregularly-spaced time series and time-to-event problems. In addition to the systematic review, the project will also document recent developments in automated data science pipelines and identify any gaps in current provision for tools that can be applied to spatio-temporal data. The project will fill any such gaps by developing initial scripts that extend existing automated data cleaning tools to add support for time-to-event analysis or temporally dependent data. The project will also conduct a limited benchmark study to evaluate the sensitivity of temporal models to different data preprocessing strategies. This study will provide insights into how the decisions made during data preprocessing can affect downstream performance of machine learning models. The ultimate goal of this project is to provide researchers with a better understanding of the best practices for cleaning, selecting, and extracting features from spatio-temporal event data. The project's results will be useful for researchers working in various fields such as epidemiology, ecology, and transportation, where spatio-temporal event data is commonly used. Coding task: Take Learn2Clean and add a new metric to it (not necessarily time-to-event; could be regression, classification etc) that isn't already implemented in the package. Then run it on one or two example datasets.

23. Simulation of crime and similar geo-spatial event data

Population-level societal events, such as civil unrest and crime, often have a significant impact on our daily life. Forecasting such events is of great importance for decisionmaking and resource allocation. Event prediction has traditionally been challenging due to the lack of knowledge regarding the true causes and underlying mechanisms of event occurrence. A recent survey looks deep learning mechansim to forecast such spatial temporal event data http://arxiv.org/abs/2112.06345. These mechansims are commonly evaluated on a handful of public data sets such as New York crime data. Assignment: The purpose of this mini project is to create a multitude of simulator that can be used to evaluate the ability to exploit complicated spatial temporal patterns to improve prediction. These events could be simulated using point processes (http://arxiv.org/abs/2011.04583), agent based models or probabilistic programming and GAN methods (https://ojs.aaai.org/index.php/AAAI/article/view/20375). The particular focus is to multi layered events of different types e.g. crime data. The burden is to show that the simulator is useful in developing models for real data.

24. Reinforcement learning for board games

This project focuses on learning strategies that help humans to improve their performance. Additionally, we will consider the problem of detecting non-human players. An initial focus will be the card game 6 nimmt!. Can we translate strategies learnt by machines into rules simple enough to be understood and used by human players? Can we distinguish 'real' players from 'bots' based on their play style? Further details are available on request. References R. Bettker, P. Minini, G. Pereira and J. V. C. Assunção, "Towards playing AIs for 7 Wonders: main patterns and strategies for 3-player games", 2021 20th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames), 2021, pp. 172-181, doi: 10.1109/SBGames54170.2021.00029.

25. Algorithmically-biased survey design

Large language models (LLMs) are powerful tools for natural language generation, but they also pose significant risks of misuse and abuse. One of the potential applications of LLMs is to design the wording of questionnaires for surveys, polls, or assessments. However, this also opens the door for generating biased or misleading questions that can manipulate the respondents' opinions, beliefs, or behaviours. This project investigates the potential severity of algorithmically-biased survey design, through carefully designed experiments under a variety of conditions and LLM frameworks.

Preferred Project:
Second Project:
Third Project:
Suggest your own topic:
Tell us why you are passionate about these topics in particular, and outline skills that you already have that are relevant for the project(s) selected.

up to 200 words ...

Submit

If you spoke to someone in our team (or someone connected to the group) already, please tell us who it was.

name, leave empty otherwise

 \wedge

 \wedge

 \wedge