XinDu

Profile Photo

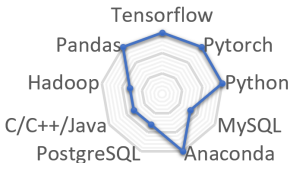
Address
PO Box 513
5600 MB Eindhoven,
the Netherlands

Tel
+44 (0)75 136 58640
+31 (0)620 843 020

Mail
x.d.du@hotmail.com

Personal Links
Homepage
Google Scholar
GitHub
Linkedin

Programming Tools



Personal Skills

Languages
Chinese ★★★★★
English ★★★★☆

Working Experience

2024-Now **Research Scientist** Great Bay University, CN
Trustworthy Machine Learning.

2021-2023 **Postdoc Research Associate** University of Edinburgh, Edinburgh, UK
Trustworthy Machine Learning.

2020-2021 **Postdoc Research Fellow** Eindhoven University of Technology, Eindhoven, Netherlands
Exceptional Model Mining, Fairness and Causal Inference

2015-2016 **Research Associate** Delft University of Technology, Delft, Netherlands
3D City Modeling.

2015 **Algorithm Research Engineer** Weibo, Beijing, China
Algorithm design and development for microblog texts labeling.

2010-2012 **Software Developer Engineer** Longmap, ShenZhen, China
4G Map Localization.

Research Experience

2017-2020 **PhD Research** Eindhoven University of Technology, The Netherlands
Department of Mathematics and Computer Science

- **Fairness in Machine Learning**. Study fairness in terms of unsupervised sensitive subgroups. Propose a fairness measure for network representation model.
- **Causal Inference**. Study the causal effect inference of individual treatments with observational data. Propose a covariate balancing technique to remove confounding bias from imbalanced data.
- **Exceptional Model Mining**. Study the exceptional multi-modal behavior on subgroups. Propose a Bayesian non-parametric model for the inference of exceptional behavior in terms of space, time and texts.
- **Learning Analytics**. Study the heterogeneous learning behavior in MOOCs.

Relevant Research Skills

- **Probabilistic Methods** Formulating problems, building statistical models and making inference from data, using point estimate or Bayesian inference.
- **Latent variable Methods** Modeling generating process and inferring parameters considering the latent variables, using EM, MCMC and variational inference.
- **Hypothesis testing** Formulating hypothesis testing with parametric/non-parametric assumptions. Validating the significance of deviations between two distributions.
- **Meta Learning** Develop learning frameworks that learns meta representation in supervised and un-supervised method.
- **Transfer Learning** Develop machine learning systems that make use of part of the learned dynamic systems that can transfer knowledge from the origin to the target environments.

- **Data mining** Designing efficient data mining tools to discover patterns in particular representations with regard to specific interestingness, e.g. exceptional performance of a model on subgroups.
- **Deep Learning Methods** Building deep neural networks and learning to optimize the performance with data. extracting useful features and applying to downstream tasks like classification and regression.
- **Causal Methods** Employing domain knowledge to build causal graph for underlying generating mechanism. Using causal graph to boost the model on tasks like counterfactual prediction, preventing the misleading of spurious associations.
- **Synthetic Analysis** Design synthetic data with specific generating process to validate particular methods.

Academic Experience

2020- **Program Committee Member**
ECML-PKDD 2020-2025, IJCAI 2021-2025, AAAI 2021-2025, UAI 2021-2025, AISTATS 2022-2025, IDA 2023, ICLR 2023, KDD 2025

2019- **Journal Reviewer** Springer
International Journal of Artificial Intelligence In Education (IJAIED)

2019 **Proceeding Chair** Würzburg, Germany
European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)

2018 **Volunteer** 's-Hertogenbosch, the Netherlands
International Symposium on Intelligent Data Analysis (IDA)

2017 **Volunteer** Eindhoven, the Netherlands
The Annual Machine Learning Conference of The Benelux (Benelearn)

2019 **Academic Visiting** Helsinki University, Finland
Data Mining Research

Awards&Copyright&Patent

2020 **Student Travel Award** New York, U.S.A
Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI)

2014 **Software Copyright:2014SR036739** Wuhan University
Multi-scale map data matching software(GeoMatching)

2014 **Software Copyright:2014SR036893** Wuhan University
POI visualization based on ubiquitous space and mapping system(POI Viewer)

2014 **Software Copyright:2014SR025346** Wuhan University
Ocean survey data management and 3D visualization information system.

2013 **Patent** CN 103473420 A
Automatic positioning method of statistical graph in zonal statistic map

Education

2017 - 2020 **PhD Student in Computer Science** Eindhoven University of Technology
*Supervisor: Prof. dr. Mykola Pechenizkiy (**m.pechenizkiy@**tue.nl),*
*Co-Supervisor : Dr. Wouter Duivesteijn (**w.duivesteijn@**tue.nl),*
Thesis: "Uncertainty in Exceptional Model Mining".

2012 - 2015 **Master's Degree in GIS** Wuhan University, China
*Supervisor : Prof. dr. Tinghua Ai (**tinghuaai@**whu.edu.cn),*
Thesis: "Information mining of place name based on crowdsourcing data".
Thesis activity carried out during the work period at Weibo, China.

2006 - 2010 **Bachelor's Degree in GIS.** Yunnan University, China
Main subjects: Mathematics, Computer Science, GIS.

Referees

Mykola Pechenizkiy
Eindhoven University of Technology *m.pechenizkiy@tue.nl*

Wouter Duivesteijn
Eindhoven University of Technology *w.duivesteijn@tue.nl*

Alexander Nikolaev
University at Buffalo-SUNY *anikolae@buffalo.edu*

Publications

X. Du, S. Yang, W. Duivesteijn, M. Pechenizkiy
Conformalized Exceptional Model Mining: Telling Where Your Model Performs (Not) Well *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2025*, to appear

X. Du, Y. Pei, W. Duivesteijn, M. Pechenizkiy
Exceptional Spatio-Temporal Behavior Mining through Bayesian Non-Parametric Modeling *Data Mining and Knowledge Discovery (ECML-PKDD Journal Track), 2020.PDF*

X. Du, Y. Pei, W. Duivesteijn, M. Pechenizkiy
Fairness in Network Representation by Latent Structural Heterogeneity in Observational Data *AAAI Conference on Artificial Intelligence (AAAI), 2020.PDF*

X. Du, L. Sun, W. Duivesteijn, A. Nikolaev and M. Pechenizkiy
Adversarial Representation Learning for Causal Effect Inference with Observational Data *Data Mining and Knowledge Discovery, 2021.PDF*

X. Du, W. Duivesteijn, M. Klabbers, M. Pechenizkiy
ELBA: Exceptional Learning Behavior Analysis *Proceedings of the Eleventh International Conference on Educational Data Mining (EDM), 2018.PDF*

X. Du, B. Legastelois, B. Ganesh, A. Rajan, H. Chockler, V. Belle, S. Anderson, S. Ramamoorthy
Vision Checklist: Testable Error Analysis of Image Models to Help System Designers Interrogate Model Capabilities *Arxiv, 2022.PDF*

X. Du, S. Ramamoorthy, W. Duivesteijn, J. Tian, M. Pechenizkiy
Beyond Discriminant Patterns: On the Robustness of Decision Rule Ensembles

Arxiv, 2021.PDF

Y. Pei, X. Du, J. Zhang, G. Fletcher, M. Pechenizkiy
struc2gauss: Structure Preserving Network Embedding via Gaussian Embedding *Data Mining and Knowledge Discovery, 2020.PDF*

Y. Pei, X. Du, J. Zhang, G. Fletcher, M. Pechenizkiy
Dynamic Network Representation Learning via Gaussian Embedding *Graph Representation Learning Workshop (NeurIPS), 2019.*

Anthony L. Corso, Sydney M. Katz, Craig Innes, Xin Du, Subramanian Ramamoorthy, Mykel J. Kochenderfer
Risk-Driven Design of Perception Systems *The 36th Conference on Neural Information Processing Systems, 2022.PDF*

SOftware

ABCEI	**Developer** *A causal inference software to estimate CATE*	Causal Inference
Annotate Optimize	**Developer** *A software to optimize map annotation*	Map Annotation