Xin Du

Phone: +44(0)7513658640
Email: mailto:x.d.du@hotmail.com
Website: https://octeufer.github.io/

WORK EXPERIENCE

POSTDOC Research Associate   **2021-2023**

- **Trustworthy Autonomous Systems.**
  Studied trustworthy autonomous systems through the lens of machine learning. Focused on the robustness of prediction or perception models against input space perturbations and datasets shifts.

  Research output: 1) Developed a tool to generate test examples based on existing datasets randomly or weakly supervised by humans to discover and report the semantic mismatch under the certain perturbations or shifts. The statistical results are organized into model cards as reports.
  2) Developed a tool to analyze the risk of decision reward under perturbations and covariate shifts.

POSTDOC Fellowship   **2020-2021**

- **Fairness, causality, and exceptional model mining.**
  Studied the application of data mining methods to discover local patterns that can be used for prediction tasks under environment changes.

  Research output: developed a tool to discovery informative local patterns and use their ensembles to perform prediction tasks, where given a causal graph indicating the underlying data generative process, the algorithm can extract constraints to regularize the predictor to ensure the highly accurate performance across different environments.

- **Fairness in Network Representation Learning**.
  Studied the concept of fairness in terms of unsupervised sensitive subgroups on network data with network representation model.

  Research output: developed an exceptional model mining method that utilized the latent representation space to induce the structural discrepancy of input network data, where conjunction and combination of attributes are reported as local patterns that indicate the meaningful subgroups compared with the original network data.

- **Conditional Average Treatment Effects.**
  Studied the problem of causal effect inference on individual level with observational data, where recent advances of representation learning methods are used to tackle the covariate shift problem between treatment and control groups.

  Research output: proposed a representation learning method that can make use of covariate information from both control and treatment groups to conduct prediction on causal responses for individuals. Adversarial learning and mutual information estimation methods are used to ensure the learning of informative representations.

- **Exceptional Model Mining on multi-modal data.**
  Studied the problem of exceptional multi-modal behavior modeling with social media data. A Bayesian non-parametric graphic model is proposed for the inference of exceptional behavior in terms of space, time, and texts.

  Research output: developed an exceptional model mining method that utilized the parametric space of the Bayesian probabilistic model to search for the exceptional behavioral patterns on social media datasets. The proposed method uses Gibbs sampling method to conduct the inference step on multi-modal social media datasets.

ACADEMIC ACTIVITIES

- **(Senior) Program Committee Member**, ECML-PKDD 2020-2023, IJCAI 2021-2023, AAAI 2021-2024, UAI 2021-2023, AISTATS 2022-2023, IDA 2023, ICLR 2023.
- **Journal Reviewer**, International Journal of Artificial Intelligence in Education (IJAIED), Data Mining and Knowledge Discovery (DAMI).
- **Proceeding Chair**, ECML-PKDD 2019.
- **Volunteer**, International Symposium on Intelligent Data Analysis (IDA), 2018.
- **Volunteer**, The Annual Machine Learning Conference of The Benelux (Benelearn), 2017.

EDUCATION

Ph.D. Computer Science, Eindhoven University of Technology, September 2020.
Thesis: "The Uncertainty in Exceptional Model Mining",
Advisor: Mykola Pechenizkiy & Wouter Duivesteijn,
https://research.tue.nl/en/publications/the-uncertainty-in-exceptional-model-mining

M.A. Geographic Information System, Wuhan University 2015.

B.S. Geographic Information System, Yunnan University 2010.

Publications

[1] 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. Work-in-progress, 2022,
https://arxiv.org/abs/2201.11674

[2] X. Du, S. Ramamoorthy, W. Duivesteijn, J. Tian, M. Pechenizkiy,
Beyond Discriminant Patterns: On the Robustness of Decision Rule Ensembles. Work-in-progress, 2021, https://arxiv.org/abs/2109.10432

[3] 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, 34, 1267-1290,
https://link.springer.com/article/10.1007/s10618-020-00674-z

[4] 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, (Vol. 34, No. 04, pp. 3809-3816),
https://ojs.aaai.org/index.php/AAAI/article/view/5792

[5] 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, 35(4), 1713-1738,
https://link.springer.com/article/10.1007/s10618-021-00759-3

[6] 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, https://eric.ed.gov/?id=ED593224

[7] *Y. Pei, X. Du, J. Zhang, G. Fletcher, M. Pechenizkiy,*
struc2gauss: Structure Preserving Network Embedding via Gaussian Embedding.
Data Mining and Knowledge Discovery, 2020,
https://link.springer.com/article/10.1007/s10618-020-00684-x

[8] Y. Wang, V. Menkovski, H. Wang, X. Du, M. Pechenizkiy,
Causal Discovery from Incomplete Data: A Deep Learning Approach. arxiv preprint, 2020,
https://arxiv.org/abs/2001.05343

[9] 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,
https://www.research.ed.ac.uk/en/publications/risk-driven-design-of-perception-systems