

# Curriculum Vitae – Wouter Duivesteijn

Date of birth        December 09, 1984  
Place of birth       Rotterdam, the Netherlands  
Nationality         Dutch  
Email address       w.duivesteijn@tue.nl  
Website             <http://wwis.win.tue.nl/~wouter/>



## Academic career

- Sep'16-today:** Assistant Professor Data Mining at Technische Universiteit Eindhoven (NL).
- Oct'15-Aug'16:** Postdoctoraal Bursaal at Universiteit Gent (BE), working on the FORSIED project: FORmalising Subjective Interestingness in Exploratory Data mining.
- Aug'15-Sep'15:** Honorary Research Associate at University of Bristol (UK), working on the FORSIED project: FORmalising Subjective Interestingness in Exploratory Data mining.
- Aug'13-Jul'15:** Wissenschaftlicher Mitarbeiter at Technische Universität Dortmund (DE), in Sonderforschungsbereich 876: Verfügbarkeit von Information durch Analyse unter Ressourcenbeschränkung (Collaborative Research Center 876: Providing Information by Resource-Constrained Data Analysis).
- Jul'09-Jul'13:** PhD candidate in the Algorithms group at Leiden Institute of Advanced Computer Science, Universiteit Leiden (NL), supervised by Joost N. Kok and Arno Knobbe (graduated Sep. 17, 2013).

## Education

- '19: University Teaching Qualification (UTQ), Technische Universiteit Eindhoven.
- '09-'13: PhD in Computer Science, Universiteit Leiden (graduated Sep. 17, 2013).
- '08-'09: MSc in History and Philosophy of Science, Universiteit Utrecht (aborted for PhD position).
- '05-'08: MSc in Applied Computing Science, Universiteit Utrecht (graduated Oct. 15, 2008).
- '05-'07: MSc in Mathematical Sciences, Universiteit Utrecht (graduated Oct. 30, 2007).
- '02-'05: BSc in Mathematics, Universiteit Utrecht (graduated Jul. 7, 2005).
- '02-'05: BSc in Computer Science, Universiteit Utrecht (graduated Jul. 7, 2005).

## Awards and funding

**Perseverance Award 2022:** the award for excellent proposals at Technische Universiteit Eindhoven's department of Mathematics and Computer Science, that narrowly didn't make the cut in competitive calls against low success rates. On the strength of a VIDI proposal that ranked 19<sup>th</sup> in a cluster where the top-17 proposals were granted, I won a Perseverance Award granting funding for a five-year PhD-TA position plus travel budget. The total funding value converts to roughly EUR 320 000.

**C.J. Kokprijs 2013:** the award for the best dissertation of the year in Universiteit Leiden's Faculty of Science. Awarded annually since 1971; I am the only computer scientist to ever win it.

Co-applicant in **EDIC: Exceptional and Deep Intelligent Coach**, a project funded in the NWO COMMIT2DATA - DATA2PERSON scheme, for a total amount of EUR 1 253 500.

**KDD 2012 student travel award:** USD 1 000 plus free registration (worth USD 380). **ICDM 2011 student travel award:** CAD 500. **ICDM 2010 student travel award:** USD 500.

# 1 Research

## 1.1 Publications

See also Google Scholar Citations: <https://scholar.google.nl/citations?user=LmsQAtAAAAAJ>.

### 1.1.1 Journal publications

14. R.M. Schouten, M.L.P. Bueno, W. Duivesteijn, M. Pechenizkiy: Mining Sequences with Exceptional Transition Behavior of Varying Order using Quality Measures based on Information-Theoretic Scoring Functions. In: *Data Mining and Knowledge Discovery* 36 (1), pp. 379–413, 2022. Impact factor: 4.8.
13. X. Du, L. Sun, W. Duivesteijn, A. Nikolaev, M. Pechenizkiy: Adversarial Balancing-based Representation Learning for Causal Effect Inference with Observational Data. In: *Data Mining and Knowledge Discovery* 35, pp. 1713–1738, 2021. Impact factor: 5.406.
12. J.M. Luna, M. Pechenizkiy, W. Duivesteijn, S. Ventura: Exceptional in So Many Ways — Discovering Descriptors that Display Exceptional Behavior on Contrasting Scenarios. In *IEEE Access* 8, pp. 200982–200994, 2020. Impact factor: 3.367.
11. W. Duivesteijn, S. Hess, X. Du: How to Cheat the Page Limit. In: *WIREs Data Mining and Knowledge Discovery* 10 (3), e1361, 2020. Impact factor: 7.250.  
Won a *Top Downloaded Article* award, for being in the top-10% most downloaded articles during its first 12 months of publication in this journal, among work published in 2019 and 2020.
10. X. Du, Y. Pei, W. Duivesteijn, M. Pechenizkiy: Exceptional Spatio-Temporal Behavior Mining through Bayesian Non-Parametric Modeling. In: *Data Mining and Knowledge Discovery* 34, pp. 1267–1290, 2020. Impact factor: 3.670.
9. C. Rebelo de Sá, W. Duivesteijn, P. Azevedo, A.M. Jorge, C. Soares, A. Knobbe: Discovering a Taste for the Unusual — Exceptional Models for Preference Mining. In: *Machine Learning* 107 (11), pp. 1775–1807, 2018. Impact factor: 2.809.
8. L. Downar, W. Duivesteijn: Exceptionally Monotone Models — the Rank Correlation Model Class for Exceptional Model Mining. In: *Knowledge and Information Systems* 51 (2), pp. 369–394, 2017. Impact factor: 2.247.
7. C. Pölitiz, W. Duivesteijn, K. Morik: Interpretable Domain Adaptation via Optimization over the Stiefel Manifold. In: *Machine Learning* 104 (2–3), pp. 315–336, 2016. Impact factor: 1.848.
6. W. Duivesteijn: Correction to Jin-Ting Zhang’s “Approximate and Asymptotic Distributions of Chi-Squared-Type Mixtures with Applications”. In: *Journal of the American Statistical Association* 111 (515), pp. 1370–1371, 2016. Impact factor: 2.016.
5. W. Duivesteijn, A.J. Feelders, A. Knobbe: Exceptional Model Mining — Supervised Descriptive Local Pattern Mining with Complex Target Concepts. In: *Data Mining and Knowledge Discovery* 30 (1), pp. 47–98, 2016. Impact factor: 3.160.
4. R.M. Konijn, W. Duivesteijn, M. Meeng, A. Knobbe: Cost-based Quality Measures in Subgroup Discovery. In: *Journal of Intelligent Information Systems* 45 (3), pp. 337–355, 2015. Impact factor: 1.000.
3. P. Lohuis, S. Faraj-Hakim, W. Duivesteijn, A. Knobbe, A.-J. Tasman: Benefits of a Short, Practical Questionnaire to Measure Subjective Perception of Nasal Appearance after Aesthetic Rhinoplasty. In: *Plastic and Reconstructive Surgery* 132 (6), pp. 913e–923e, 2013. Impact factor: 3.328.

2. P.J.F.M. Lohuis, S. Hakim, A. Knobbe, W. Duivesteijn, G.M. Bran: Split hump technique for reduction of the overprojected nasal dorsum – a statistical analysis on subjective body image in relation to nasal appearance and nasal patency in 97 aesthetic rhinoplasty patients. In: *Archives of Facial Plastic Surgery*, 2012, 14 (5), pp. 346–353. Impact factor: 1.463.
1. S. Hakim, A. Knobbe, W. Duivesteijn, P.J.F.M. Lohuis: Results of a screening questionnaire measuring physical perception of patients undergoing esthetic rhinoplasty: a statistical analysis. In: *Nederlands Tijdschrift voor Keel-Neus-Oorheelkunde* (Dutch Journal for Otorhinolaryngology), 2010 (2), p. 100. Impact factor: 0.

### 1.1.2 Conference publications

25. R.F.A. Verhaegh, J.J.E. Kiezebrink, F. Nusteling, A.W.A. Rio, M.B. Bendiscek, W. Duivesteijn, R.M. Schouten: A Clustering-inspired Quality Measure for Exceptional Preferences Mining — Design Choices and Consequences. In: *Proceedings of the 25<sup>th</sup> International Conference on Discovery Science (DS 2022)*, pp. 429–444.  
Acceptance rate: 0.4655 (27 out of 58). Including short papers: 0.6724 (39 out of 58).
24. J.F. van der Haar, S.C. Nagelkerken, I.G. Smit, K. van Straaten, J.A. Tack, R.M. Schouten, W. Duivesteijn: Efficient Subgroup Discovery Through Auto-Encoding. In: *Proceedings of the 20<sup>th</sup> International Symposium on Intelligent Data Analysis (IDA 2022)*, pp. 327–340, 2022.  
Acceptance rate: 0.4366 (31 out of 71).
23. R.M. Schouten, W. Duivesteijn, M. Pechenizkiy: Exceptional Model Mining for Repeated Cross-Sectional Data (EMM-RCS). In: *Proceedings of the 2022 SIAM International Conference on Data Mining (SDM 2022)*, pp. 585–593.  
Acceptance rate: 0.2785 (83 out of 298).  
Supplementary material is available on Figshare: <https://doi.org/10.6084/m9.figshare.18688220>
22. Y. Soons, R. Dijkman, M. Jilderda, W. Duivesteijn: Predicting Remaining Useful Life with Similarity-based Priors. In: *Proceedings of the 18<sup>th</sup> International Symposium on Intelligent Data Analysis (IDA 2020)*, pp. 483–495, 2020. Acceptance rate: 0.3947 (45 out of 114).
21. X. Du, Y. Pei, W. Duivesteijn, M. Pechenizkiy: Fairness in Network Representation by Latent Structural Heterogeneity in Observational Data. In *Proceedings of the 34<sup>th</sup> AAAI Conference on Artificial Intelligence (AAAI 2020)*, pp. 3809–3816, 2020.  
Acceptance rate: 0.2056 (1591 out of 7737).
20. A. Belfodil, W. Duivesteijn, M. Plantevit, S. Cazalens, P. Lamarre: DEvIANT: Discovering significant exceptional (dis-)agreement within groups. In: *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2019)*, pp. 3-20, 2019. Acceptance rate: 0.1771 (130 out of 734).
19. S. Hess, W. Duivesteijn: k is the Magic Number — Inferring the Number of Clusters Through Nonparametric Concentration Inequalities. In: *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2019)*, pp. 257–273, 2019. Acceptance rate: 0.1771 (130 out of 734).
18. S. Hess, W. Duivesteijn, P. Honysz, K. Morik: The SpectACl of Nonconvex Clustering: A Spectral Approach to Density-Based Clustering. In: *Proceedings of the 33<sup>rd</sup> AAAI Conference on Artificial Intelligence (AAAI 2019)*, pp. 3788–3795, 2019. Acceptance rate: 0.1621 (1150 out of 7095).
17. X. Du, W. Duivesteijn, M. Klabbers, M. Pechenizkiy: ELBA: Exceptional Learning Behavior Analysis. In: *Proceedings of the 11<sup>th</sup> International Conference on Educational Data Mining (EDM 2018)*, pp. 312–318, 2018. Acceptance rate: 0.4207 (61 out of 145).

16. J. Lijffijt, B. Kang, W. Duivesteijn, K. Puolamäki, E. Oikarinen, T. De Bie: Subjectively Interesting Subgroup Discovery on Real-valued Targets. In: Proceedings of the 34<sup>th</sup> IEEE International Conference on Data Engineering (ICDE 2018), pp. 1352-1355, 2018.  
Acceptance rate: 0.2311 (98 out of 424).  
A substantially longer version appeared on arXiv: <https://arxiv.org/abs/1710.04521>
15. W. Duivesteijn, T. Farzami, T. Putman, E. Peer, H.J.P. Weerts, J.N. Adeggest, G. Foks, M. Pechenizkiy: Have It Both Ways — from A/B Testing to A&B Testing with Exceptional Model Mining. In: Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2017), part III, pp. 114–126, 2017.  
Acceptance rate: 0.2903 (27 out of 93).
14. C. Rebelo de Sá, W. Duivesteijn, C. Soares, A. Knobbe: Exceptional Preferences Mining. In: Proceedings of the 19<sup>th</sup> International Conference on Discovery Science (DS 2016), pp. 3–18, 2016.  
Acceptance rate: 0.5000 (30 out of 60).
13. L. Downar, W. Duivesteijn: Exceptionally Monotone Models — the Rank Correlation Model Class for Exceptional Model Mining. In: Proceedings of the 15<sup>th</sup> IEEE International Conference on Data Mining (ICDM 2015), pp. 111–120, 2015.  
Acceptance rate: 0.0843 (68 out of 807). Including short papers: 0.1821 (147 out of 807).  
Distilled from Lennart's Bachelor thesis, available at:  
[http://sfb876.tu-dortmund.de/PublicPublicationFiles/downar\\_2014a.pdf](http://sfb876.tu-dortmund.de/PublicPublicationFiles/downar_2014a.pdf)
12. W. Duivesteijn, J. Thaele: Understanding Where Your Classifier Does (Not) Work. In: Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2015) (III), pp. 250–253, 2015.  
Acceptance rate: 0.4828 (14 out of 29).
11. W. Duivesteijn, J. Thaele: Understanding Where Your Classifier Does (Not) Work — the SCaPE Model Class for EMM. In: Proceedings of the 14<sup>th</sup> IEEE International Conference on Data Mining (ICDM 2014), pp. 809–814, 2014.  
Acceptance rate: 0.1953 (142 out of 727).  
A substantially longer version appeared as Technical Report of the SFB 876 at the TU Dortmund:  
[http://sfb876.tu-dortmund.de/PublicPublicationFiles/duivesteijn\\_thaele\\_2014a.pdf](http://sfb876.tu-dortmund.de/PublicPublicationFiles/duivesteijn_thaele_2014a.pdf)
10. J. Witteveen, W. Duivesteijn, A. Knobbe, P. Grünwald: RealKRIMP – Finding Hyperintervals that Compress with MDL for Real-Valued Data. In: Proceedings of the 13<sup>th</sup> International Symposium on Intelligent Data Analysis (IDA 2014), pp. 368–379, 2014.  
Acceptance rate: 0.4800 (36 out of 75).  
Distilled from Jouke's Bachelor thesis, available at:  
<http://liacs.leidenuniv.nl/assets/Bachelorscripties/2012-13JoukeWitteveen.pdf>
9. M. Meeng, W. Duivesteijn, A. Knobbe: ROCsearch – An ROC-guided Search Strategy for Subgroup Discovery. In: Proceedings of the 2014 SIAM International Conference on Data Mining (SDM 2014), pp. 704–712, 2014.  
Acceptance rate: 0.3084 (120 out of 389).
8. W. Duivesteijn, A. Knobbe: Exceptional Model Mining – Describing Deviations in Datasets. In: Proceedings of the 22<sup>nd</sup> Belgian-Dutch Conference on Machine Learning (BENELEARN 2013), p. 86, 2013.  
Acceptance rate: 0.8919 (33 out of 37).
7. R.M. Konijn, W. Duivesteijn, W. Kowalczyk, A. Knobbe: Discovering Local Subgroups, with an Application to Fraud Detection. In: Proceedings of the 17<sup>th</sup> Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2013), pp. 1–12, 2013.  
Acceptance rate: 0.1134 (39 out of 344). Including short presentations: 0.2849 (98 out of 344).

6. W. Duivesteijn, E. Loza Mencía, J. Fürnkranz, A. Knobbe: Multi-label LeGo – Enhancing Multi-label Classifiers with Local Patterns. In: Proceedings of *the 11<sup>th</sup> International Symposium on Intelligent Data Analysis (IDA 2012)*, pp. 114–125, 2012.  
Acceptance rate: 0.2250 (18 out of 80). Including poster presentations: 0.4375 (35 out of 80). A substantially longer version appeared as Technical Report TUD-KE-2012-02 of the TU Darmstadt: <http://www.ke.tu-darmstadt.de/publications/reports/tud-ke-2012-02.pdf>
5. G. Ribeiro, W. Duivesteijn, C. Soares, A. Knobbe: Multilayer Perceptron for Label Ranking. In: Proceedings of *the 22<sup>nd</sup> International Conference on Artificial Neural Networks (ICANN 2012)*, pp. 25–32, 2012.  
Acceptance rate: 0.6559 (162 out of 247).
4. W. Duivesteijn, A. Feelders, A. Knobbe: Different Slopes for Different Folks – Mining for Exceptional Regression Models with Cook’s Distance. In: Proceedings of *the 18<sup>th</sup> ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2012)*, pp. 868–876, 2012.  
Acceptance rate: 0.1762 (133 out of 755).
3. W. Duivesteijn, A. Knobbe: Exploiting False Discoveries – Statistical Validation of Patterns and Quality Measures in Subgroup Discovery. In: Proceedings of *the 11<sup>th</sup> IEEE International Conference on Data Mining (ICDM 2011)*, pp. 151–160, 2011.  
Acceptance rate: 0.1285 (101 out of 786). Including short papers: 0.1883 (148 out of 786).
2. W. Duivesteijn, A. Knobbe, A. Feelders, M. van Leeuwen: Subgroup Discovery meets Bayesian networks – an Exceptional Model Mining approach. In: Proceedings of *the 10<sup>th</sup> IEEE International Conference on Data Mining (ICDM 2010)*, pp. 158–167, 2010.  
Acceptance rate: 0.0903 (72 out of 797). Including short papers: 0.1945 (155 out of 797).
1. W. Duivesteijn, A. Feelders: Nearest Neighbour Classification with Monotonicity Constraints. In: Proceedings of *the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2008) (I)*, pp. 301–316, 2008.  
Acceptance rate: 0.1919 (100 out of 521).

### 1.1.3 Workshop publications

7. W. Duivesteijn, T.C. van Dijk: Exceptional Gestalt Mining: Combining Magic Cards to Make Complex Coalitions Thrive. In: *the 8<sup>th</sup> Workshop on Machine Learning and Data Mining for Sports Analytics*, pp. 191–204, 2021.
6. S.B. van der Zon, W. Duivesteijn, W. van Ipenburg, J. Veldsink, M. Pechenizkiy: ICIE 1.0: A Novel Tool for Interactive Contextual Interaction Explanations. In: *MIDAS 2018, PAP 2018: ECML PKDD 2018 Workshops*, pp. 81–94, 2019.
5. S. van der Zon, O. Zeev Ben Mordehai, T. Vrijdag, W. van Ipenburg, J. Veldsink, W. Duivesteijn, M. Pechenizkiy: BoostEMM — Transparent Boosting using Exceptional Model Mining. In: Proceedings of *the Second Workshop on Mining Data for financial applications (MIDAS 2017)*, pp. 5–14, 2017.
4. W. Duivesteijn, M. Meeng, A. Knobbe: ROCsearch in a Wider Context – A ROC-Guided Search Strategy for Subgroup Discovery and Beyond. In: Proceedings of *the First International Workshop on Learning over Multiple Contexts (LMCE 2014)*, 2014.
3. W. Duivesteijn: A Short Survey of Exceptional Model Mining – Exploring Unusual Interactions Between Multiple Targets. In: Proceedings of *the 2014 International Workshop on Multi-Target Prediction (MTP 2014)*, 2014.
2. M. Meeng, W. Duivesteijn, A. Knobbe: ROCsearch – An ROC-Guided Search Strategy for Subgroup Discovery. In: Proceedings of *the 2014 Workshop on Knowledge Discovery, Data Mining and Machine Learning (KDML 2014)*, p. 180, 2014.

1. R.M. Konijn, W. Duivesteijn, M. Meeng, A. Knobbe: Cost-based Quality Measures in Subgroup Discovery. In: Proceedings of the 3<sup>rd</sup> *Quality Issues, Measures of Interestingness, and Evaluation of data mining models workshop (QIMIE 2013)*, PAKDD Workshops, pp. 404–415, 2013.

## 1.2 Editor of books

- \* W. Duivesteijn, A. Siebes, A. Ukkonen (eds.): *Advances in Intelligent Data Analysis XVII*; 17th International Symposium, IDA 2018, 's-Hertogenbosch, the Netherlands, October 24–26, 2018, Proceedings. Part of the *Lecture Notes in Computer Science* book series (LNCS, volume 11191), Springer, Cham.
- \* M. Atzmueller, W. Duivesteijn (eds.): *Artificial Intelligence*; 30th Benelux Conference, BNAIC 2018, 's-Hertogenbosch, the Netherlands, November 8–9, 2018, Revised Selected Papers. Part of the *Communications in Computer and Information Science* book series (CCIS, volume 1021), Springer, Cham.

## 1.3 Co-promotor of PhD students

Feb.'20-today: Rianne M. Schouten

Jan.'17-Feb.'21: Simon B. van der Zon

Jan.'17-Sep.'20: Xin Du (defended thesis *The Uncertainty in Exceptional Model Mining* Sep. 28, 2020)

Jan.'17-Nov.'19: Negar Ahmadi (defended thesis *EEG Microstate and Functional Brain Network Features for Classification of Epilepsy and PNES* Nov. 19, 2019)

## 1.4 Member of doctoral committees

Apr.'23: Juan Antonio Bellido Jiménez, Universidad de Córdoba (ES)

Feb.'23: Lu Yin, Technische Universiteit Eindhoven

Jan.'23: Loong Kuan Lee, Monash University (AU)

Feb.'20: Yulong Pei, Technische Universiteit Eindhoven

Nov.'19: Negar Ahmadi, Technische Universiteit Eindhoven

Oct.'18: Jianpeng Zhang, Technische Universiteit Eindhoven

Sep.'17: Firat Ismailoglu, Universiteit Maastricht

Dec.'16: Claudio Rebelo de Sá, Universiteit Leiden

## 1.5 Keynote speaker

Oct.'16: keynote “Some Recent Advances in Exceptional Model Mining — Unusual Preferences and more” at the PEPS Préfute workshop, IRISA INRIA Rennes, France. This was a quarterly workshop organized by and for a consortium of collaborating French data mining research groups, who invited international experts to deliver the keynote talks.

## 1.6 Organizer of scientific meetings

**Local Chair** of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence (UAI 2022) in Eindhoven, the Netherlands, together with Sibylle Hess.

**Proceedings Chair** of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2019) in Würzburg, Germany, together with Sibylle Hess and Xin Du. Reprised role without co-chairs at ECML PKDD 2020 in Gent, Belgium. Reprised role together with Sibylle Hess at ECML PKDD 2022 in Grenoble, France. Role to be reprised at ECML PKDD 2024 in Vilnius, Lithuania.

**General chair** of the Thirtieth Benelux Conference on Artificial Intelligence (BNAIC 2018) in 's-Hertogenbosch, the Netherlands, together with Martin Atzmueller.

**General chair** of the Seventeenth International Symposium on Intelligent Data Analysis (IDA 2018) in 's-Hertogenbosch, the Netherlands.

**Conference chair** of the annual machine learning conference of the Benelux (Benelearn 2017) in Eindhoven, the Netherlands, together with Mykola Pechenizkiy.

**Workshop chair** of Silver 2012 (together with Joaquin Vanschoren), collocated with ECML PKDD in Bristol, UK. This workshop was dedicated to learning from unexpected results, and disseminating the resulting knowledge.

## 1.7 Member of scientific societies

'18-now: member of the IDA Advisory Board (<http://www.ida-society.org/advisoryboard.php>).

'19-'20: member of the Association for the Advancement of Artificial Intelligence (AAAI).

## 1.8 Reviewing activities

**Reviewer of grant proposals** for for the United States – Israel Binational Science Foundation (BSF) and the Deutsche Forschungsgemeinschaft (DFG).

**Member of the Editorial Board** of *Data Mining and Knowledge Discovery*.

**Member of the Guest Editorial Board** of the ECML PKDD journal track 2020–2022, implemented in partnership with the Machine Learning Journal (MACH) and the Data Mining and Knowledge Discovery Journal (DAMI).

**Member of the scientific committee** for the books: *Data Science for Healthcare: Methodologies and Application*, published by Springer Heidelberg in 2019, and *Data Science for Economics and Finance: Methodologies and Applications*, published by Springer in 2021.

**Senior PC member** of IDA 2020, 2022, and 2023.

**Program committee member** of AAAI 2020–2021, DS 2019–2023, ECML PKDD 2015–2019 and 2021–2023, ICDM 2014–2017 and 2019, IDA 2014–2017, IDEA@KDD 2015–2017, IDM 2022, IJCAI 2016, ISMIS 2017, KDD 2015–2016 and 2018–2019, KDID 2022, KDML 2014–2015, LMCE 2015, MIDAS 2019–2023, MLSA 2022–2023 SDM 2018–2024, UAI 2021.

**Best reviewer award** at ECML PKDD 2021.

**Reviewer for journals** DAMI, IJBDE, INS, JMLR, KAIS, MACH, NCAA, NEUCOM, PR, TKDD, TKDE.

## 1.9 Research visits

**Oct.'12:** For one day, visited the Varieng Institute at the University of Helsinki, Finland. Collaborated with Prof. Terttu Nevalainen, Tanja Säily, and Mikko Hakala on subgroup discovery and data cleaning in the Parsed Corpus of Early English Correspondence.

**Feb.'12:** For two days, visited the Integrated Data Mining group at the Fraunhofer IAIS, Bonn, Germany. Exchanged ideas with Dr. Stefan Rüping, Dr. Henrik Großkreutz, and Dr. Mario Boley; collaborated with the latter and Sandy Moens on integrating coupling-from-the-past-sampling with subgroup discovery.

**Nov.'11-Dec.'11:** For two weeks, visited the Department of Criminology at Simon Fraser University, Burnaby, Canada. Collaborated with Dr. Richard Frank on improving a webcrawler built to detect child pornography.

**Oct.'10-Nov.'10:** For two weeks, visited the Knowledge Engineering group at the Technische Universität Darmstadt, Germany. Collaborated with Eneldo Loza Mencía and Prof. Johannes Fürnkranz on integrating multi-label classification with exceptional model mining having deviating Bayesian network structure as target.



## 2 Teaching

Obtained University Teaching Qualification (BKO/UTQ) certificate, the standard Dutch certification documenting my ability to teach in higher education, from the Technische Universiteit Eindhoven on May 14, 2019.

Developed from scratch the new course **JM0150: Data Mining**, in the Data Science and Entrepreneurship master at JADS. First lecture was on the morning of Tuesday, September 13, 2016; i.e.: on my ninth working day at TU/e. The course got rave reviews in the student evaluations.

### 2.1 Courses taught

For JM0150 and 2IMM20/2AMM20, I was the responsible lecturer.

#### 2.1.1 Academic year '23/'24

**2AMM20:** Research Topics in Data Mining; 5 ECTS course in the DSAI/CSE master at TU/e; 115 students.

**BSc HA:** Co-coordinator of the AI track in the Honors Academy at TU/e; 25 students.

#### 2.1.2 Academic year '22/'23

**2AMM20:** Research Topics in Data Mining; 5 ECTS course in the DSAI/CSE master at TU/e; 110 students.

Received an *Excellent Course Evaluation* award (“Pluim”) from the TU/e CSE program management, for obtaining excellent student evaluations.

**BSc HA:** Co-coordinator of the AI track in the Honors Academy at TU/e; 22 students.

#### 2.1.3 Academic year '21/'22

**2AMM20:** Research Topics in Data Mining; 5 ECTS course in the DSAI/CSE master at TU/e; 195 students (same course code as previous year, new content; instead of a foundational course, we now presuppose knowledge of basic DM/ML concepts, and we demand that all the students push the boundaries of scientific knowledge in their final research project).

Received an *Excellent Course Evaluation* award (“Pluim”) from the TU/e CSE program management, for obtaining excellent student evaluations.

#### 2.1.4 Academic year '20/'21

**2AMM20:** Foundations of Data Mining; 5 ECTS course in the Computer Science and Engineering master at TU/e; 305 students (new course code, same content as previous year; course taught fully online due to COVID-19).

**BSc HA:** Coordinator of the AI track in the Honors Academy at TU/e; 22 students.

#### 2.1.5 Academic year '19/'20

**2IMM20:** Foundations of Data Mining; 5 ECTS course in the Computer Science and Engineering master at TU/e; 310 students.

**2IMM00:** Seminar Data Mining; 5 ECTS course in the Computer Science and Engineering master at TU/e, preparing 59 students for their MSc thesis in the data mining group.

**2IX30:** Responsible Data Science; 5 ECTS course in the Computer Science bachelors at TU/e; 55 students.

**BSc HA:** Midway through the year, took over the Track Coordinator task of the AI track in the Honors Academy at TU/e; 15 students.

### 2.1.6 Academic year '18/'19

**2IMM20:** Foundations of Data Mining; 5 ECTS course in the Computer Science and Engineering master at TU/e; 134 students.

**2IID0:** Web Analytics; 5 ECTS course in the Computer Science bachelors at TU/e; 121 students.

**2IAB0:** Data Analytics for Engineers; 5 ECTS course in the Bachelor College, compulsory for all 2023 freshman Bachelor students TU/e-wide.

### 2.1.7 Academic year '17/'18

**JM0150:** Data Mining; 6 ECTS course in the Data Science and Entrepreneurship master at JADS; 34 students.

**2IID0:** Web Analytics; 5 ECTS course in the Computer Science bachelors at TU/e; 97 students.

**2IAB0:** Data Analytics for Engineers; 5 ECTS course in the Bachelor College, compulsory for all 2300 freshman Bachelor students TU/e-wide.

### 2.1.8 Academic year '16/'17

**JM0150:** Data Mining; 6 ECTS course in the Data Science and Entrepreneurship master at JADS; 16 students.

**PDEDMM:** Data Mining Module; course in the Data Science PDEng at JADS; 9 students.

**2IID0:** Web Analytics; 5 ECTS course in the Computer Science bachelors at TU/e; 77 students.

## 2.2 Invited lectures

**Jan.'23:** guest lecture *How to Lie with Data*; part of the bachelor course “Data Mining and Machine Learning” (2IIG0), TU/e.

**Jan.'22:** guest lecture *How to Lie with Data*; part of the bachelor course “Data Mining and Machine Learning” (2IIG0), TU/e.

**Dec.'21:** invited talk *Exceptional Model Mining — Finding Interpretable Subgroups Exhibiting Exceptional Behavior*; opening lecture of the SIKS (Netherlands research school for Information and Knowledge Systems) course “Foundations of Data Science: Data Mining”.

**Mar.'16:** guest lecture *Exceptional Model Mining — Understanding Unusual Interactions Between Multiple Targets*; part of the master course “Advanced Data Mining”, Universiteit Antwerpen, Belgium.

## 2.3 Students supervised

33. Isaam Al Habash ('23)
32. Iko Vloothuis ('23)
31. Pim Rietjens ('23)
30. Victoria Tascau ('23, DEPAR)
29. Niels Schelleman ('23)
28. Bart Engelen ('23, PricewaterhouseCoopers)
27. Mika van Loon ('22)
26. Márton Bendegúz Bendicsek ('22)
25. Qiang Fang ('20/1, Nationale Politie)
24. Cees Gniewyk ('20/1, Siemens Technology München)
23. Nikki Branderhorst ('20/1)
22. Noud Frints ('20/1)
21. Wout de Ruiter ('18/21)
20. Andrea Benevenuta ('20, IKNL)
19. Armand Duijn ('20, StudyPortals)
18. Jelle van der Ster ('20)
17. Shama Khalil ('20)
16. Natarajan Chidambaram ('19/20, Philips)
15. Vignesh Srinivasan ('19/20, Philips)
14. Chiara Attanasio ('19, IKNL)
13. Mathyn Scheerder ('18/9, Nationale Politie)
12. Brent van Strien ('18/9)
11. Charlotte Peperkamp ('18/9, Politie Amsterdam)
10. Keje Sinnige ('18/9, Eneco Group)
9. Lucas Otten ('18, Shareship)
8. Akarsh Sinha ('18, Vencomatic)
7. Youri Soons ('18, Sitech)
6. Joery de Vos ('17/8)
5. Maurice Houben ('16/7, Simbuka)
4. Boy Raaijmakers ('16/7)
3. Ahmet Celikkaya ('15)
2. Lennart Downar ('14)
1. Jouke Witteveen ('12)

The projects of students 1–29 are completed; the projects of students 30–33 are ongoing.

Student 1 was from Universiteit Leiden, students 2 and 3 were from Technische Universität Dortmund, students 9–11 were from JADS; the rest is from TU/e.

Students 1, 2, and 4 did their Bachelor thesis under my supervision, students 5 and 16 did an internship under my supervision, student 6 had me supervising his Honors Master program, student 8 did his PDEng thesis with me, student 26 followed a Capita Selecta under my supervision. The other 25 students on this list have/had me as their MSc thesis supervisor.

## 2.4 Teaching assistant/tutor

**Spring '16:** Waarschijnlijkheidsrekening en Statistiek.

**Summer '15:** Wissensentdeckung in Datenbanken.

**Winter '14/5:** Software-Praktikum.

**Summer '14:** Software-Praktikum.

**Winter '13/4:** Software-Praktikum.

**Fall '12:** Fundamentele Informatica voor Informatica & Economie.

**Spring '12:** Fundamentele Informatica 3.

**Fall '11:** Fundamentele Informatica 2.

**Spring '11:** Fundamentele Informatica 3.

**Fall '10:** Fundamentele Informatica 2.

**Spring '10:** Algoritmiëk.

**Fall '09:** Concepten van Programmeertalen.

**Fall '08:** Wiskundige Technieken voor Fysici.

**Spring '08:** Wiskunde voor Chemici.

**Fall '07:** Wiskunde voor Natuurwetenschap & Innovatiëmanagement, Wiskunde voor Chemici.

**Fall '06:** Lineaire Algebra, Wiskundetoepassingen.

**Fall '05:** Lineaire Algebra.

**Fall '04:** Lineaire Algebra, Wiskunde voor Informatici, Logica en Verzamelingen.

**Spring '04:** Wiskunde voor Biologen.

### 3 Committees and outreach, both within and outside of academia

I was invited to represent data mining on an episode of the TU Eindhoven podcast **Sound of Science**, where the Flemish cabaretier Lieven Scheire has a ~ 45min one-on-one discussion with a scientist, for general audience consumption. Recorded in October 2019; appeared online in February 2020 as episode 16. All episodes of the podcast can be found at <https://www.tue.nl/en/research/sound-of-science/>. My episode can be listened to at <https://soundcloud.com/user-563469139/sound-of-science-16-wouter>

**May '20:** day chair of EIASI-affiliated symposium “Decision making with AI in industry”; cancelled due to COVID-19.

**Nov.'19:** as part of my DSCE task package, I was **day chair** at the Data Science Summit, a ~ 400 person event at the Muziekgebouw Eindhoven, on November 12, 2019.

#### 3.1 Invited talks

**Sep.'20:** invited talk at ABN AMRO’s financial crime week.

**Feb.'19:** invited talk “On the TU/e Data Mining Group, Exceptional Model Mining, and the Opel Meriva” at the first Inspiration Session of the data science innovation team, know your customer department of ABN AMRO.

**Nov.'18:** invited talk “Proeftuin data-analyse en AI. Leren over verandergedrag van criminelen door verschillende lokale dataprojecten aan elkaar te koppelen” as half of a TU Eindhoven double bill with Marius Monen, at the Jamsessie Trends of the FIOD (“Fiscale Inlichtingen- en OpsporingsDienst”; the anti-fraud agency of the Dutch tax authorities).

#### 3.2 Institutional responsibilities

**Aug.'22-Sep.'25:** deputy member of the Portfolio Committee UTQ at the Department of Mathematics and Computer Science of TU/e.

**Apr.'18-Dec.'20:** research program manager for Customer Journey and Responsible Data Science research programs at the Data Science Center Eindhoven.

Functioned as **external committee member** in the TU/e thesis committees of (all are MSc theses, except for Rens’; his was a BSC thesis):

- |                                  |                               |
|----------------------------------|-------------------------------|
| 11. Maiko Bergman ('23)          | 5. Nena O'Driscoll ('19, NS)  |
| 10. Andrei Rykov ('23)           | 4. Anja Syring ('18, PulseDC) |
| 9. Thomas Quadt ('23)            | 3. Jeroen Hoogers ('18)       |
| 8. Mathijs Boezer ('22)          | 2. Li Wang ('17)              |
| 7. Nandini Rajasekar ('21)       | 1. Bob Giesbers ('16)         |
| 6. Rens Oostenbach ('20, Tiqets) |                               |

In '16/7, I spent about 30 hours working as a **scientific consultant for Witteveen & Bos**, who investigated property damage due to the Earth tremors in the Northern Dutch province of Groningen. These damages could be caused by the gas extraction taking place in this province, or by other causes. The consultancy consisted of exploring whether Exceptional Model Mining with regression as model class, could find coherent subgroups of the property dataset where the amount of damages related unusually to input parameters.

'16-today: faculty member at Technische Universiteit Eindhoven.

'07-'08: chairman of the board of students' association A-Eskwadraat, for Mathematics, Physics, Computer Science, and Information Science, with more than 1700 members.

'06-'07: member of the "faculteitsraad Bètawetenschappen" (science faculty council) and the "departementsraad Wiskunde" (mathematics department council) at Universiteit Utrecht.

'05-'06: member of the Mathematics department's board of education at Universiteit Utrecht.

'04-'05: member of the "opleidingscommissie Wiskunde" (mathematics education council) at Universiteit Utrecht.