Main research interest

The chair studies data mining (DM) techniques and knowledge discovery approaches that are at the core of data science. The group is known for its contributions to the areas of predictive analytics, automation of machine learning and networked science, subgroup discovery and exceptional model mining, and similarity computations on complex data. Its research is inspired by theoretical computer science, systems development and real-world applications of (big) data-driven discovery in healthcare, banking, energy, retail, telecom, and education among others. We develop generic approaches and specialized techniques that cover a wide range of descriptive, predictive and prescriptive analytics and work effectively with text, image, transactional, graph and time-series data in a responsible manner. E.g. we use Deep Learning methods to develop models for high dimensional heterogeneous, unstructured and evolving data and apply this models to areas such as medical imaging, genomics, anomaly detection and sentiment analysis. We further work on methods for analyzing and explaining the model’s decisions and performance and facilitate effective DM with domain expert in the loop.

The group is actively involved in DCS/e, leading the Customer Journey and taking part in Smart Manufacturing and Maintenance research programs and collaborates with Data Engineering, Visualization, Security, HCI, and Process mining research groups.

Scientific staff

Prof. Mykola Pechenizkiy (head of the group)
Predictive analytics, Evolving data streams, Handling concept drift, Complex networks, Responsible data science

Dr. Anne Driemel
High-dimensional computational geometry, Clustering and classification in non-Euclidean spaces, and Information retrieval for time series and trajectories data

Dr. Wouter Duivesteijn
Supervised learning, Subgroup discovery, and Exceptional model mining

Dr. Vlado Menkovski
Machine learning, Deep learning, Representation learning, and Reinforcement learning

Dr. Joaquin Vanschoren
Automation of machine learning, Meta-learning, and Deep learning

Furthermore 9 PhDs and 1 PD are working on various Data Science projects.

Success stories

We have created OpenML: an online collaborative platform for studying machine learning techniques. OpenML is used by almost 2,000 researchers, students, and practitioners world-wide, and contains around 20,000 datasets, 3,000 machine learning workflows, and 1.7 million shared experiments. It has won the Dutch Data Prize, as well as backing from Microsoft Research. It is crucial for the development of automated machine learning that is adopted by companies such as Philips.

Further information at www.OpenML.org

Project examples

- **NWO RATE-Analytics** *(with Tilburg University, Rabobank and Achmea)* “Next generation predictive analytics for data-driven banking and insurance”.

- **Impulse KYC-Analytics** *(with Rabobank)* “Know your customer predictive analytics” project aims at developing approaches for effective DM on heterogeneous and evolving data sources with expert-in-the-loop.

- **STW CAPA** *(with Adversitement and StudyPortals)* “Context-aware predictive analytics” advanced the current state of the art in Web analytics.

- **NWO Veni** “Detection methods for similarity structures in time-dependent data” develops foundations for advanced time series and trajectories clustering.

- **H2020 SODA** *(ICT-2016-1; Big Data PPP)* “Scalable Oblivious Data Analytics” facilitates secure DM; together with Crypto group we develop practical approaches for DM with multi-party computation.
Data science is an interdisciplinary field that uses a variety of techniques to create value based on extracting knowledge and insights from available data. Data science is applied everywhere: in business, health, industry, finance, government, education, and also in scientific research.

The Data Science Center Eindhoven (DSC/e) is TU/e’s response to these challenges and possibilities. By bringing top scientists and students from over thirty research groups from different TU/e departments together on specific topics, we can tackle the most challenging scientific and societal challenges. All involved groups made a one-page description of their main research interests and the involved staff with their key expertise, like the one you’re holding now.

**Mathematics and Computer Science**
- Algorithms
- Applied Geometric Algorithms
- Architecture of Information Systems
- Data Mining
- Mathematical Image Analysis
- Probability
- Security of Embedded Systems
- Software Engineering & Technology
- Statistics
- Stochastic Operations Research
- System Architecture & Networking
- Visualization
- Web Engineering

**Electrical Engineering**
- Cognitive Internet of Things
- Control Systems
- Electrical Energy Systems
- Signal Processing Systems

**Built Environment**
- Building Lighting
- Information Systems in the Built Environment
- Real Estate Management & Urban Planning
- Urbanism and Urban Architecture

**Industrial Engineering & Innovation Sciences**
- Human Technology Interaction
- Information Systems
- Innovation, Technology Entrepreneurship & Marketing
- OPAC: Freight Transport & Logistics
- OPAC: Maintenance & Manufacturing
- OPAC: Supply chain management
- Philosophy & Ethics

**Biomedical Engineering**
- Cardiovascular Biomechanics
- Computational Biology
- Medical Image Analysis