Madelon Hulsebos

Amsterdam, Netherlands | +31 6 83 23 37 67 | m.hulsebos@uva.nl | madelonhulsebos.github.io

EDUCATION

University of Amsterdam – Amsterdam, Netherlands
PhD, Computer Science *2020 – present*

Delft Institute of Technology – Delft, Netherlands
MSc, Computer Science *2016 – 2018*
BSc, Technology, Policy and Management *2011 – 2015*

ACADEMIC EXPERIENCE

University of Amsterdam – Amsterdam, Netherlands
PhD Researcher, INDE Lab *2020 – present*
- Developed a research agenda for improving table understanding systems in practice, proposing 4 research projects around methods, a dataset and system integrations in downstream applications.
- Executed 3 out of 4 proposed projects, partially at Sigma Computing.

Massachusetts Institute of Technology – Cambridge, Massachusetts, United States
Visiting Collaborator, MIT Media Lab *2018 – 2019*
- Led and maintained a research project on semantic table understanding.
- Contributed to a data visualization training and benchmarking repository.

Delft Institute of Technology – Delft, Netherlands
Graduate Teaching Assistant, Pattern Recognition & Web Information Systems groups *2017 – 2018*
- Supported 250+ graduate students throughout projects and tutorials of multiple courses.
- Evaluated student assignments, projects and presentations.

Aalto University – Helsinki, Finland
Research and Teaching Assistant, Machine Learning for Big Data *2017*
- Designed the assignments of an ML course with 500+ students.
- Built research experiments for semi-supervised learning over networks.

INDUSTRY EXPERIENCE

Sigma Computing – San Francisco, California, United States
Research Intern *2021 – present*
- Developed a framework and system for deploying semantic table understanding systems in practice.
- Contributed to an interactive business analysis system.
- Published the resulting papers at the Conference of Innovative Data Systems Research.

KPN, HEINEKEN – Rotterdam, Amsterdam
Data Scientist *2019 – 2021*
- Developed and integrated an ML system for financial forecasting that produced more accurate forecasts.
- Built ML products for evaluating marketing campaigns.
- Mentored several junior data scientists and analysts.
- Initiated and led reading groups and implemented scrum processes.
- Gave presentations and tutorials on data validation, software practices and transfer learning.

ACADEMIC SERVICE

NeurIPS '21, EBDT '22
Reviewer *2021 – present*
- Reviewed contributions for the Datasets & Benchmarks track of NeurIPS.
- Reviewed contributions for the Industrial track of the Extended Database Technology conference (EDBT).

SemTab challenge '21
Organizing Committee Member *2021*
- Compiling a benchmark dataset from GitTables.
- Supporting and reviewing submissions for the column type annotation task.

LEADERSHIP

UniPartners Delft – Delft, Netherlands
Supervisory Board Member *2017 – present*
- Supervise the financial and strategic positions of the company.
- Implemented structured financial control, KPI oriented leadership and a supervision cycle.

Executive Board Member *2015 – 2016*
- Controlled and optimized the quality of internal and external processes.
- Daily management of projects, contributing to a revenue of over €100 K.
- Moderated the CRM system (Salesforce) and trained 10+ boards.

SELECTED PUBLICATIONS

Hulsebos, M., Gathani, S., Gale, J., Dillig, I., Groth, P., Demiralp, C. "Making Table Understanding Work in Practice", arXiv preprint arXiv:2109.05173 (2021).

Hulsebos, M., Demiralp, C, Groth, P. "GitTables: A Large-Scale Corpus of Relational Tables", arXiv preprint arXiv:2106.0725 (2021).

Zhang, D., Suhara, Y., Li, J., **Hulsebos, M.**, Demiralp, C., Tan, W. "Sato: Contextual semantic type detection in tables", in Proceedings of the VLDB Endowment, 2020.

Hulsebos, M., Hu, K., Bakker, M., Zgraggen, E., Satyanarayan, A., Kraska, T., Demiralp, C., Hidalgo, C. "Sherlock: A deep learning approach to semantic data type detection", in ACM SIGKDD. ACM, 2019.

Hu, K., Gaikwad, N., **Hulsebos, M.**, Bakker, M., Zgraggen, E., Hidalgo, C., Kraska, T., Li, G, Satyanarayan, A., Demiralp, C. (2018) "VizNet: Towards a large-scale visualization learning and benchmarking repository", in ACM CHI. ACM, 2019.

EXPERTISE

Research: table understanding, knowledge bases, natural language processing, weak supervision, deep learning.
Data science: machine learning, causal inference, bayesian modeling, data validation.
Tools: Scikit-learn, TensorFlow, Keras, PyStan, NLTK, Stanford CoreNLP, Great Expectations, React, Git, Airflow.
Leadership: mentoring, consulting, workshop organization, Scrum, Agile.
Programming Languages: Python, Java, Javascript, Matlab, R, Visual Basic, LaTeX.
Natural Languages: Dutch (native), English (proficient), German (intermediate).