## CURRICULUM VITAE

**Marcel H. Schubert /**
01.08.1992

**https://github.com/mhschubert/Portfolio**

**SUMMARY**

Data Scientist trained in Economics and Computer Science with 4-year experience in applying machine learning to questions in social and economic sciences. Proficient in implementing end-to-end project pipelines and working with a wide range of data types in different projects, such as NLP classifiers for political ideology predictions on court decisions or defining a behavior type space using unsupervised learning on time-series data. Seeking to leverage my skills in statistical analysis and machine learning to deliver action-oriented solutions to complex business problems.

**EDUCATION**

09/2017 – 11/2021 **Ph.D. Economics**
*University of Cologne*
Thesis:     Behavioral Economics & Machine Learning

10/2017 – 11/2019 **M.Sc. Cyber Security**
University of Lancaster

10/2015 – 09/2017 **M.Sc. in Economics**          **B.Sc. in Computer Science**
University of Bonn              University of Bonn

10/2012 – 02/2015 **B.Sc. in Economics**
Ludwig-Maximilians University (LMU Munich)

**PRACTICAL EXPERIENCE**

10/2017 – today **Research Fellow**
*Max Planck Institute for Research on Collective Goods*
- Researching the application of machine learning models for behavioral economics
- Working on end-to-end project pipeline, from scraping to prediction
- Designing, programming, and statistical analysis of experiments and behavioral studies
- Visualizing data insights and storytelling for scientific publication
- Analyzing data and hypotheses using statistics

09/2017 – 11/2021 **Ph.D. Projects**
- Defining behavioral arch-types in group interactions using unsupervised learning on time-series data
- Black-box bias analysis on the impact of machine learning model design for the COMPAS algorithm (predictions for bailing and jailing decisions)
- Stability and relevance analysis on feature weight matrix of classifiers in NLP authorship analysis tasks using a large-scale experiment
- Classifier for the political ideology of US court opinions/texts and extension of labeled texts database from 20,000 to 1 mil

03/2019 – 05/2019 **Guest Researcher**
*GESS ETH Zurich Working Group for Law, Economics & Data Science*
- Classifier System for Legal Texts (>1000 characters)

05/2015 – 07/2015 **Internship Consultancy**
*Helbling Business Advisors Stuttgart*

## QUALIFICATIONS

**VOLUNTARY WORK**
Ph.D. Representative
*Max Planck Institute for Research on Collective Goods*
Financial Officer of University Group
*RCDS Bonn*
Member of the Public Relations Committee of the Student Representation
*AStA University of Bonn*

**LANGUAGES**
German - Native
English - Fluent
Spanish - Basic
Swahili - Basic

**EXPERTISE**
Statistical Analysis & Econometrics (Proficient)
Machine Learning [classical & recent models] (Proficient)
Natural Language Processing [NLP] (Proficient)
Data Cleaning & Pre-Processing (Proficient)
Experimental Design & Causal Inference (Competent)
Predictive Analysis (Competent)
Biologically Inspired Computing (Competent)

**PROGRAMMING**
Python (Proficient)
R (Proficient)
JavaScript (Competent)
C (Basic)
Java (Basic)

**DATA SCIENCE & ML**
Data Science Toolkits [e.g., Pandas, SciPy, NumPy] (Proficient)
Scikit-learn (Proficient)
Spacy (Proficient)
PyTorch (Competent)
TensorFlow / Keras (Basic)

**DATABASE TECHNOLOGIES**
SQL (Competent)
MongoDB (Basic)

**OTHER**
LaTeX (Proficient)
Linux – Shell/Bash (Proficient)
Version Control / Git (Proficient)
Slurm & HPC Computing (Proficient)
HTML/CSS (Competent)
oTree (Competent)