Joschka Braun

joschkacbraun@gmail.com · (+49)175 8901489 · Website · **Linked**in

EDUCATION

University of Tübingen   Tübingen, Germany
Master of Science, Machine Learning   2022–2024
   – *Current Grade average:* 1.4

University of Tübingen   Tübingen, Germany
Bachelor of Science, Computer Science   2020–2022
   – *Final Grade:* 1.3 (Top 5% of class)
   – *Thesis:* Verbal Epistemic Uncertainty Estimation for Numeric Values with GPT-3

University of Tübingen   Tübingen, Germany
Mathematics   2017–2020
   – Completed lectures in Linear Algebra, Analysis, Measure Theory, Stochastics, Numerical Analysis, Algebra, ...

WORK EXPERIENCE

KASL - Krueger AI Safety Lab   Cambridge, United Kingdom
Full-time, paid Research Internship   Jun 2024 - Dec 2024
   – Researching why steering vectors are unreliable and what their limitations are, supervised by Dmitrii Krasheninnikov and David Krueger.

Health NLP Lab   Tübingen, Germany
Part-time, Student Researcher   Oct 2023 - Jun 2024
   – Research in controlled text generation and topical summarization, under Ali Bahrainian and Carsten Eickhoff.

RAWLAB GmBH   Reutlingen, Germany
Student Employee, Machine Learning Engineer   May 2021–Jan 2023
   – Implemented timeseries analysis algorithms, trained supervised ML models such as LinearSVC, and applied transfer learning techniques on MobileNetV2 to classify spectrograms of time series measurements.

AI SAFETY EXPERIENCE

ASET - AI Safety Engineering Taskforce   Remote
Participant   Sep 23rd – Nov 1st, 2024
   – As part of the Benchmarks team of ASET, I'm integrating the BIG-Bench Hard benchmark into the inspect AI framework of the UK AI Safety Institute.

Cambridge ERA:AI Fellowship   Cambridge, United Kingdom
Participant   Jul 1st – Aug 23rd, 2024
   – As a technical AI Safety fellow I did research on the limitations of contrastive activation steering in LLMs and attended various events learning more about how to mitigate risks from frontier AI.

REFERENCES

Dmitrii Krasheninnikov: Supervisor of research project at KASL
David Krueger: Senior Supervisor of research project at KASL
Seyed Ali Bahrainian: Supervisor of research project at Health NLP group
Marius Hobbhahn: Supervisor of Bachelor's Thesis
Contact details available on their websites or upon request.

SKILLS

Programming: Python, PyTorch, TensorFlow, Transformers (Huggingface), NumPy, Pandas, Scikit-Learn, Keras
Relevant University Courses: Deep Learning, Reinforcement Learning, Natural Language Processing, Statistical Machine Learning, Probabilistic Machine Learning, Self-Driving Cars, Linear Algebra
Linguistic Proficiency: German (Native), English (Proficient)

Miscellaneous

Scholarship: Bertha von Suttner-Studienwerk (2022–2024), awarded to 15 students out of 170 applicants, based on their academic achievements, social engagement, and commitment to humanist values.

Debating: Two-time winner of the German Debating Championship (DDM) 2019 & 2020, won the largest and most prestigious German-speaking debate competition for two consecutive years.