Jaco du Toit

Cape Town, South Africa | +27 78 834 4936
jacowdutoit11@gmail.com
github | linkedin

SUMMARY

I am a computer scientist and researcher with a strong foundation in machine learning, natural language processing (NLP), and large language models. My academic background and professional experience have enabled me to work on a range of projects, from building NLP systems to developing machine learning pipelines for real-world applications. I am excited by the rapid advancements in AI systems, but concerned about the catastrophic risks they may pose. Therefore, I am passionate about exploring AI safety, interpretability, and alignment in my future research to ensure that AI benefits all humanity.

EDUCATION

STELLENBOSCH UNIVERSITY   *Jan 2023 - March 2024*

MSc COMPUTER SCIENCE
• Thesis: Hierarchical Text Classification with Transformer-based Language Models.
Cum Laude (85%).

STELLENBOSCH UNIVERSITY   *Jan 2022 - Dec 2022*

BSc (HONS) COMPUTER SCIENCE
• Focus on machine learning with mini thesis in machine learning and natural language processing.
Cum Laude (86%).

STELLENBOSCH UNIVERSITY   *Jan 2019 - Dec 2021*

BCOM MATHEMATICAL SCIENCE (DATA SCIENCE)
• Double major in Computer Science and Operations Research.
Cum Laude (87%) with 90% aggregate for final year.

EXPERIENCE

MATOGEN APPLIED INSIGHTS   *May 2024 - Present*

DATA SCIENTIST.
• Designed and implemented data pipelines to support machine learning workflows.
• Developed machine learning models for predictive analytics, optimizing performance and scalability.
• Built NLP solutions, including chatbots, tailored to meet specific client needs.

STELLENBOSCH UNIVERSITY   *2022*

COMPUTER SCIENCE TEACHING ASSISTANT.

SKILLS

PROGRAMMING LANGUAGES	***Experienced:*** Python	***Familiar:*** C	Java	R	Javascript	SQL
SOFTWARE DEVELOPMENT	GIT	Docker	CLI	Agile Methodology
FRAMEWORKS & LIBRARIES	PyTorch	Tensorflow	HuggingFace	Pandas	PySpark	Scikit-learn
LANGUAGES	***Fluent:*** English	Afrikaans	***Beginner:*** Dutch	German	Spanish

PUBLICATIONS

• du Toit, J., Dunaiski, M. (2023). Hierarchical Text Classification Using Language Models with Global Label-Wise Attention Mechanisms. In: Artificial Intelligence Research. SACAIR 2023. Communications in Computer and Information Science, vol 1976. Springer, Cham. https://doi.org/10.1007/978-3-031-49002-6_18
• du Toit, J., Dunaiski, M. (2024). Prompt Tuning Discriminative Language Models for Hierarchical Text Classification. Natural Language Processing. In press. doi:10.1017/nlp.2024.51

AWARDS AND HONOURS

STELLENBOSCH UNIVERSITY   *2022*

RECTOR'S AWARD FOR ACADEMIC EXCELLENCE.

STELLENBOSCH UNIVERSITY   *2023*

FACULTY OF SCIENCE POSTGRADUATE RESEARCH CONFERENCE: BEST ORAL PRESENTATION.