

Jaco DU TOIT  
Cape Town, South Africa | +27 78 834 4936  
jacowdutoit11@gmail.com  
[github](#) | [linkedin](#)

## SUMMARY

I am a highly motivated computer scientist specialising in data science, machine learning, and natural language processing. Over the past two years I have immersed myself in intensive research, solidifying my expertise in natural language processing and forming a keen interest in deep learning. My interests extend to various domains of machine learning, including computer vision, speech, time series forecasting, and reinforcement learning. Most of all, I am driven by an enthusiasm for leveraging machine learning to solve interesting and impactful real-world problems. The combination of my interest in the practical application and theoretical exploration of machine learning allows me to bring a comprehensive perspective to any project.

## 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.

### STELLENBOSCH UNIVERSITY

2023

CO-SUPERVISOR FOR COMPUTER SCIENCE MASTERS STUDENT.

### STELLENBOSCH UNIVERSITY

2022

COMPUTER SCIENCE TEACHING ASSISTANT.

## SKILLS

PROGRAMMING LANGUAGES	<b>Experienced:</b> Python   <b>Familiar:</b> Java   C   R   Javascript   SQL
SOFTWARE DEVELOPMENT	GIT   Docker   CLI   Agile Methodology
FRAMEWORKS & LIBRARIES	PyTorch   Tensorflow   HuggingFace   Pandas   PySpark   Scikit-learn
LANGUAGE MODELS	<b>Experienced:</b> BERT   RoBERTa   ELECTRA   <b>Familiar:</b> GPT   Llama
LANGUAGES	<b>Fluent:</b> English   Afrikaans   <b>Beginner:</b> 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](https://doi.org/10.1007/978-3-031-49002-6_18)

## AWARDS AND HONOURS

### STELLENBOSCH UNIVERSITY

2022

RECTOR'S AWARD FOR ACADEMIC EXCELLENCE. AWARD FOR BEING ONE OF THE TOP ACHIEVERS AT THE UNIVERSITY.