ARLO HEINRICH STEYN

ABOUT ME

As a dedicated 4th-year data science student with a strong background in computer science, I bring a unique blend of advanced mathematical rigor, statistical modeling expertise, and programming proficiency. My degree has provided me with a deep understanding of core mathematical concepts such as linear algebra, calculus, probability, and optimization, alongside advanced statistical methods like hypothesis testing, predictive modeling, and stochastic processes. This solid foundation equips me to solve complex, highly technical challenges at the intersection of data science and computer science. Beyond academics, I am committed to continuous growth and improvement. I embrace challenges with enthusiasm and am dedicated to refining both my technical skills and my ability to communicate complex findings effectively.

CONTACT

PHONE: +27 71 286 8232

EMAIL: arlo.steyn@gmail.com

HOBBIES

Sports
General exercise
Gaming
Learning
Reading

TECHNICAL SKILLS

Programming: Python, Java, C/C++, R, JavaScript, SQL, TypeScript, Bash

Data Science: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, Keras

Machine Learning:
-**Supervised:** Proficient in regression, classification, ensemble methods etc, and deep learning (FNNs, CNNs, RNNs).
-**Unsupervised:** Skilled in clustering (K-Means, DBSCAN), dimensionality reduction (PCA, t-SNE), and anomaly detection.

NLP: BERT, LLaMA, GPT models, Word2Vec, Topic Modeling, Sentiment Analysis, Spacy

Web Development: React.js, Node.js, Flask, HTML/CSS, FastAPI

Databases: PostgreSQL, MySQL, Azure-hosted databases

Midstream College (IEB)
Matric - 2020
Subjects (English, Afrikaans, Mathematics, Life Orientation, Information Technology, Tourism, Business Studies)
6 Distinctions, 84% Avg.

Stellenbosch University
Student
February 2021 to current (Fourth year)
Course studying: **Bachelor of Data Science (BDatSci) -Focal Area Computer Science**.
Graduating end 2025.

The Code Academy
Professional Certification: Full-Stack Engineer
Completed 2023
Exams passed (Web Development Foundations, Building Interactive Websites, Front-End Development, Back-End Development, Full-Stack Development

PROJECTS

Compiler Project
Built a compiler from scratch in c (Using bytecode from the JVM), using a context free grammar called ALAN.

File Compression Program
Coded a file compression system using Intel IA-32 Bit Assembly.

File Sharing and Transferring Web App
Built a live file transferring and sharing react web application using typescript with node.js as backend and a hosted postgre sql azure database.

Time Tracking Web App
Built a time tracking web app using Flask in python with a hosted postegre sql azure database.

Web Scraper for git API
Built a web scraper for the git API in python using Flask.

Classification Model Analysis
Performed statistical analyses on a large dataset to predict if a customer will make a purchase based on which advert was used, in python (anaconda) using KNN, Random Forest and XGBoost.

Lung Cancer Prediction Model
Developed and optimized machine learning models (Logistic Regression with Lasso, Ridge, Elastic Net regularization, and Random Forest Classifiers) to predict lung cancer risk (low, medium, high). Applied advanced EDA, hyperparameter tuning, and cross-validation techniques to improve model accuracy and reduce false positives/negatives.

LLM Bot for Q&A
Fine tuned a Llama 3.2 8 billion parameter model using cloud computing resources with scraped data (50 000 Q&A Pairs) from 'Cross Validated' to create a Q&A bot for Data Science, Artificial Intelligence and Machine Learning.