SHANE VAN HEERDEN

MACHINE LEARNING ENGINEER

+27 76 812 4903
shaneavh8@gmail.com
shaneandrewvanheerden
Remote / Cape Town, South Africa

Profile

Highly motivated Machine Learning Engineer with experience in developing AI solutions for real-world problems. Strong knowledge in Deep Learning, Python, and Natural Language Processing, and a passion for discovering insights from structured and unstructured data. Experienced in leading data science projects and delivering productizable AI solutions.

Experience

Luno

Sep 2021 – Present

Data Scientist

- Developed real-time customer AML & fraud risk models, lead development of model evaluation feedback loop, built PySpark-based Databricks pipelines and Looker dashboards for model monitoring, and advised on appropriate trade-offs between model precision and recall.
- Data Science Lead for Zendesk CRM migration project, identified critical data sources to ingest based on reporting requirements, coordinated with Engineers on appropriate CRM configurations, and advised on a new data architecture.
- Produced accurate inbound customer message forecasts for workforce hiring and scheduling optimisation, explored correlation between Bitcoin spending behaviour and inbound messages volumes, and benchmarked predictive performance of ARIMA and Prophet forecasting models.
- Automated various productivity reports for Luno's Customer Service Workforce Team, developed PySpark-based Databricks pipelines, and displayed productivity metrics to individual Customer Success Team Leaders via Looker dashboards.

Cape AI

Mar 2020 – Aug 2021

Machine Learning Engineer

- Matched resumes to job specifications using advanced NLP techniques, developed production-ready preprocessing pipelines, stored and queried results from an Elasticsearch database, fine-tuned BERT-based models for text classification and NER, and served final results via interactive Streamlit dashboard.
- Connected employees for knowledge-sharing opportunities by analyzing unstructured text communications, leveraged state-of-the-art NLP techniques, built production-ready data pipelines interfacing with a Neo4j graph database, and deployed and monitoring an online learning model recommending possible connection opportunities.
- Advised a client's product review website on decreasing webpage creation time by employing automatic content generation techniques, optimised for SEO objectives, and developed a proof-of-concept pipeline using Python and Huggingface Transformers.
- Developed a machine learning solution for identifying sensitive client data fields in a client's data exchange platform, employed Logistic Regression classifier and established a feedback loop between the client's local application and Azure cloud platform.

DataProphet

Jan 2020 – Mar 2020

Data Science Intern

- Projects involved discovering good operating regions for clients' manufacturing machinery by employing fast Fourier transforms together with variational autoencoders using the Scikit-learn and Tensorflow Python packages.

Education

PhD (Data Science)
2017 – 2019

Stellenbosch University

- Research focused on the design and development of a Data Mining framework for quantifying and characterising road accident risk using machine learning and implemented as a Python-based solution.

BEng (Industrial Engineering)
2013 – 2016

Stellenbosch University

- Invited to the Golden Key International Honour Society.
- Won the prize for the best computer-based decision support system in the final year project.

Skills

Data Science

Natural Language Processing, Data Analysis & Visualisation, Supervised/Unsupervised/ Reinforcement Learning and Computer Vision

Programming languages

Python, R, Bash, MATLAB/Octave, C, C#, HTML, VBA, LaTeX, Cypher and SQL

Tools & Databases

Git, Docker, GIS, Neo4j, Elasticsearch and PostgreSQL

Libraries & Frameworks

Numpy, Pandas, Matplotlib, Plotly, SpaCy, Tensorflow, Pytorch, PySpark, Scikit-learn, Transformers, SHAP, BeautifulSoup, Selenium, Flask, Streamlit, Logging and Pytest

Platforms

Databricks, Looker, GitLab, AWS, Azure, Slack, Miro, Google Suite

Methodologies

Agile, Scrum, Operations Research and Simulation

Academic exposure

European Conference on Operational Research (EURO) *Jun 2019*

University College Dublin, Ireland

- Presented research findings at the 30th EURO conference in Dublin, exposed to new ideas in the fields of Operations Research and Machine Learning.

Deep Learning Indaba *Sep 2018*

Stellenbosch University, South Africa

- Attended a week-long meeting of the African Machine Learning community, exposed to teaching, research, exchange, and debate around the state-of-the-art in Machine Learning and Artificial Intelligence.

Operations Research Society of South Africa Conference *2016, 2017, 2018 & 2019*

South Africa

- Presented research findings at the 2016, 2017, 2018 & 2019 ORSSA annual conferences.

References

Contact details available on request