Robert J Czanik, Ph.D.

ML Engineer

robert@skyworkz.nl

I am an experienced machine learning engineer who has worked on numerous data science solutions for multiple companies and clients. My main focus is providing end-to-end cloud-based solutions to clients. This all is done by doing all the work needed for conceptualizing, developing, and productization of a machine learning solution. I also have the certifications to back up my skillset: MLOps, data science, data engineering, and full-stack cloud development. My background is in open-source technologies, mainly on the Linux-tech stack. Of these technologies I the am strongest in RDMS, Python, R, ETL and the development of data science use cases as web-apps. My most recent experience lies in working as an MLOps engineer, where data science use cases are developed to completion and put into a production state. I am also a strong leader with a strategic and entrepreneurial mindset; Enabling me to have also led teams in building solutions using Agile and Scrum. As a strong communicator I had no challenge in regular engagements with stakeholders, technical teams, and the delivery team. Which also requires me to listen to all parties involved and to understand the problem holisticaly. This is done by creating a lateral understanding of a problem and develop the best outcome, and elegant solution for any client.

SKILLS

AZURE	SYNAPSE, DEV OPS, DATA FACTORY, SQL SERVER, AKS, AZURE FUNCTIONS
AWS	RED SHIFT, LAMBDA, EC2, S3
IBM CLOUD	OPENSHIFT, DEV OPS, FUNCTIONS, WATSON STUDIO, NLU
ML SUITES	TENSORFLOW, PYTORCH, NUMPY, SCIPY, PYSPARK
VISUALIZATION	MATPLOTLIB, SEABORN, TABLEAU, GGPLOT2
ML METHODS	TIME SERIES, DEEP LEARNING, CLUSTERING, COMPUTER VISION, NLP, REGRESSION, GENAI, LLM
DATABASES	ELASTIC SEARCH, SQL SERVER, POSTGRESQL, DATABRICKS, APACHE SPARK
CONTAINERS	VAGRANT, KUBERNETES, DOCKER, KEDA.SH
METHODOLOGIES	AGILE, DEVOPS, MLOPS
INFRA AS CODE	ANSIBLE
PROGRAMMING	PYTHON, BASH, TYPESCRIPT, JAVASCRIPT, JAVA, SCALA, POWERSHELL
SOFT SKILLS	LEADERSHIP, STRATEGIC THINKING, SOLUTION ARCHITECTING, ENTERPRENEURIAL

EMPLOYMENT

Cloud/Data Engineer, Skyworkz

Utrecht, Netherlands — 2023-01 — Present

Skyworkz is a boutique cloud consultancy based in Utrecht and helps companies in their Cloud journey. In my role as a machine learning engineer I help companies by building data science focused solutions primarily in the cloud. Working for Skyworkz gives me an significant advantage: I can use the the experience of brilliant colleagues as reference over a large range of technologies. This enables me to develop the highest quality solutions for clients and keep improving delivery.

Senior Machine Learning Engineer, Union Investment (via Skyworkz)

Remote, client in Germany — 2023-02 — 2023-09

For the period spent at Union-Investment I worked as a staff engineering that supported a larger team in the design and development of a machine learning based stock price validation system from various stock price providers. The system worked on a principle of streaming prices received from various stock brokers over an Azure service bus through a verification model. The price validator would receive feedback from end-user stock brokers that would indicate if prices validated by the model were correct. The retraining, serving and deployment of the model was automated. For a newly trained model to have been deployed to production a performance report was sent with a Jira ticket to the business department to accpet or reject.

Highlights:

- Doing the solutions architecture of the retraining of a new model and the deployment of the model to production.
- Developing KEDA-based services that provide model predictions via streaming records to and from Azure service bus.
- Develop a service that deploys the latest version of a model to production or rollback to a previous version used in production on demand.

Azure DevOps | MLflow | Azure Machine Learning Studio | Keda | Azure Service Bus | SQL Server | Event based architecture

Machine Learning engineer, ABN Amro (via Capgemini)

2022-09 — 2022-12

Utrecht, Netherlands

I worked as a part of ABN AMRO ML-Engineering's Centre of Excellence, which is a central team serving the whole bank's data science engineering needs. Across the bank there are various data science teams who develop use cases that serve internal or clients with improving value. The engineering part of the journey is all run on the Azure stack, where airflow is used to run the workflows for data science use cases seemlessly and MLflow is used to track and give feedback to data science teams on model performance.

Highlights:

- Taking various new use cases from either POC or MVP state to production ready.
- Serving as a single point of contact for very specialised data science teams their engineering needs during deployment.

Azure Dev Ops | MLflow | Airflow | Python | Databricks

TAM+Machine Learning engineer, Logius (via Capgemini)

2022-04 — 2022-08

Den Haag, Netherlands

Logius run specific systems for the 'Ministerie van Binnenlandse Zaken' on a virtual cloud-based in AWS. To monitor the performance of the system the ELK stack is used for logging data. To determine when the system might become heavy under load demand a toy model was built to show whether the system could fall over in periods of high demand, and plan upscaling the system before such events occurred. This was done using the ELK stack's machine-learning libraries.

Highlights:

- Supported the infastructure and server runtime on the servers running kubernetes clusters
- Developed a toy model that makes recommendations for system configurations based on performance demand.

ELK | RHEL | Ansible | AWS

Machine Learning engineer, Nationale Nederlanden (via Capgemini)

Den Haag, Netherlands

2022-01 — 2022-03

I designed and developed a time series based POC solution for Nationale Nederlanden, which predicted the number of offices occupiers to their main offices. This was a time series based regression model predicting the maximum number of offices workers expected in their office facilities.

Highlights:

- The POC model was up to 78% accurate during corona times and was put into production.

Databricks | Azure Data Factory

Data engineer, MWingz (via Capgemini)

Den Haag, Netherlands | 2021-11 — 2021-12

In a project with a Belgium telecoms company, Mwingz I worked on the design and development of extracting metadata from legacy XML files to a Dynamics 365 data warehouse on Azure.

Highlights:

- The ETL operation significantly increased the transfer time fot the business from the legacy system.

Databricks | Azure Data Factory | Dynamics 365 database | XML

Senior Consultant | Cloud DevOps/Machine Learning engineer, Capgemini

Utrecht, Netherlands | 2021-11 — 2022-12

On internal projects, I developed an NLP sentiment analysis model for healthcare providers based on the Dutch website `zorgkaartnederland.nl`, where subjects that are a concern in the healthcare system by the larger public were highlighted. Followed by a project for the internal subsidiary team where the ratio of flora against paving is identified from optical and infrared images, which would help trace possible causes for flooding events during extreme weather events.

Highlights:

- Lead two junior data scientist in developing part of a computer vision idetifying objects in gardens model using satelite images.
- Developed a POC that highlighted subjects/areas of the Dutch healthcare system that could be improved on based on online healthcare providors.

TensorFlow | NLP | Satelite Imaging | Sentiment Analysis

Technical Lead, EXPLORE-AI

Durban, South Africa (Remote) | 2021-03 — 2021-10

I lead two teams building analytics tools for a large water utility company in the UK. the first data product generated reports that were sent to environmental groups and regulators. The second product generated analytics reports on both clean and wastewater asset reporting of alarms across the whole network of the water utility.

Highlights:

- Designed and lead a team of six developing one MVP to completion, in time and in budget.
- Designed and lead a team of eight while training a new technical lead up the a point shortly before completion of another MVP.

Databricks | Azure data factory | React | SQL Server

Senior Data Scientist, Explore-AI

Durban, South Africa (Remote) | 2020-10 — 2021-02

I helped a data engineering team maintain our Databricks environment on the Azure cloud platform for a large water utility company. I also looked into time series based datasets from which asset failures were supposed to identify failing assets.

Highlights:

- Spontanious time serious analysis of assets highlighted when assets most likely failed
- Helped with the upgrade and testing of Databricks environment across the whole organisation to a new LTS version.

Databricks | PySpark | Times series

Data Scientist/Analyst, ThoroughTec Simulation

Durban, South Africa | 2019-11 — 2020-10

ThoroughTech simulation is a company that builds simulators of large land vehicles used in the mining industry. As part of their product line of providing driver training, and online training management system (TMS) was planned and developed. As part of this training product and IoT device monitoring, driver behaviour would be developed. This device would be attached to a truck and monitor if someone drove an unsafe manner.

Highlights:

- Primarily wrote SQL stored procedures used for aggregation and data analytics
- Containerised the solution

SQL Server | Synthetic datasets | Times series | Docker | IoT

Data Scientist, Vantage Data

Durban, South Africa | 2019-03 — 2019-08

I worked on various time-series analysis in Alteryx, Python and R. The main focus was to model demand forecasting figure for a major retail supplier of fast-moving consumer goods.

Highlights:

- The model, which was a hybrid model of an XGBOOST regression and LSTM deep learning model which was accurate up to 90% for their 12-month and 52-week prediction figures

Time Series | Deep learning | Predictive Analytics

Data Engineer, Blake Group

Durban, South Africa | 2018-10 — 2019-03

The Blake Group is primarily a debt collection company that also hosts some digital products. One of these products, WiFire provides free WiFi access in large shopping centers. The primary selling point of the WiFi service is targeted online advertising to users of the WiFi network, based on identifying user interests.

Highlights:

- I optimize and maintain ETL data pipelines running data on AWS EC2 VMs. The pipelines transmuted the URLs visited by each user of the WiFi network and matched them with previous instances of the same users on the network.
- I developed a POC which identifies a user's interests based on the content of unencrypted URLs that they visited.
- The POC application was written in the form of a Python web scraper together with an interest builder running a NLP model.

EDUCATION

North-West University

2013-12 — 2019-12

Doctor of Philosophy - PhD: Space Physics

North-West University

2011-12 — 2012-12

MSc:Astrophysics

North-West University

2009-12 — 2010-12

BSc. Hons: Physics

North-West University

2006-12 — 2008-12

BSc: Physics-Mathematics and Applied mathematics

CERTIFICATION

Machine Learning Engineering for Production (MLOps) Specialization, Coursera/Deeplearining.ai

2022

Java Programming and Software Engineering Fundamentals Specialization, Coursera/Duke University

2022

Applied Cloud Development Specialization, Coursera/IBM 2022

IBM Full Stack Software Developer Specialization, Coursera 2022

Cloud Application Development Foundations Specialization, Coursera/IBM

2021

Advanced Data Science with IBM Specialization, Coursera/IBM 2021

DeepLearning.AI TensorFlow Developer Specialization, Coursera/Deeplearining.ai

2020

IBM Data Science Specialization, Coursera/IBM 2020

IBM AI Engineering Specialization, Coursera 2020

LANGUAGES

ENGLISH	NATIVE SPEAKER
DUTCH	INTERMEDIATE
AFRIKAANS	NATIVE SPEAKER