Norbert Eke
Data Scientist

Recently moved to Zürich with EU citizenship

📱 +1 250 878 7547 | ✉ norberteke@protonmail.com | 🏠 norberte.github.io | GitHub | 🔗 LinkedIn

Education

Master of Computer Science - Specialization in Data Science   *Ottawa, ON, Canada*
Carleton University   *Sept 2018 - June 2020*
Graduated with a GPA of 11.4/ 12. Master's Thesis in NLP, Machine Learning, and Data Mining

Bachelor of Science - Honours in Computer Science, Minor in Data Science   *Kelowna, BC, Canada*
University of British Columbia   *Sept 2014 - June 2018*
Graduated with a GPA of 3.8/ 4.0. Honours Thesis in NLP and Machine Learning

Experience

Data Scientist   *Ottawa, ON, Canada*
Health Canada   *February 2021 - present*
- Developing **NLP & predictive modeling** for Health Canada's **Risk Intelligence and Strategic Operations** unit
- Responsible for researching & implementing text processing, **keyword extraction**, and predictive modeling using **ML**
- Working in an **agile** framework with non-technical people **to gather requirements** and to **design innovative and complex techniques** for automating workflows, ensuring customer satisfaction, usability, and effectiveness

Graduate Research & Teaching Assistant   *Ottawa, ON, Canada*
Carleton University   *Sept 2018 - April 2020*
- Hosted office hours & graded assignments for the **Introduction to Databases** and **Distributed Computing** courses
- Conducted research on *Cross-Platform Software Developer Expertise Learning* in Prof. Baysal's *Software Analytics* lab

Data Scientist Intern   *Ottawa, ON, Canada*
National Research Council Canada (NRC)   *May 2019 - August 2019*
- Successfully delivered a fully working novel solution for a government client with NRC's Data Analytics Centre
- Performed various **time series analysis** of textual data using **Bayesian changepoint detection** algorithms
- Delivered an extensive report on the analysis results, presented findings to the client, and shipped the source code

Research Assistant   *Kelowna, B.C, Canada*
University of British Columbia   *May 2017 - August 2017*
- Conducted research on **feature-based opinion mining** and delivered a presentation about our findings – GitHub
- Developed automated opinion detection from reviews about product features using **Machine Learning** techniques

Skills

Programming	Python, R, SQL, Java,HTML, Algorithm Design, Data Structures, Object Oriented Programming
Software Engineering	Experienced with **Client-Server Applications** (projects), **Agile Development**, Algorithm Implementation, CI & CD, **Git**, Deployment, **Unit Testing, Writing Reports & Documentation**
Databases	**Extensive experience with SQL**, Database Architecture Design, **MySQL**, PostgreSQL, NoSQL
Cloud & Deployment	Azure, Flask, Docker, Tableau, Cognos, Colaboratory, Familiarity with Linux environments
ML & DS Libraries	**TensorFlow**, **Keras**, **Scikit-Learn**, Imbalanced-Learn, **Tidyverse**, Caret, PyCaret, StatsModels, Seaborn, **Numpy, Pandas, Matplotlib**, SciPy, Scikit-Optimize, SpaCy, **Gensim, NLTK**, dedupe
Machine Learning	Traditional ML (Random Forest, Decision Trees, SVM), **Regression**, **Classification**, **Clustering**, Time-series Forecasting, **Feature Selection**, **Outlier Detection**, NLP, Hyper-parameter Optim.
Deep Learning	Neural Nets for **NLP** and **Computer Vision** Applications using MLP, CNN, RNN, LSTM, GRU, GAN
Statistics	Descriptive Stats, Probability Theory, **Bayesian Statistics**, **Statistical Modeling**, Time Series Analysis, Statistical Language Models, Inference, Dimensionality Reduction, Hypothesis Testing
Data Analysis	Extensive experience with **Data Cleaning**, **Wrangling**, **Visualization** (Tableau, d3.js, ggplot2), EDA, **Feature Engineering**, **Data Mining**, **Predictive Modeling**, Handling **Unstructured Data**
Interpersonal Skills	**Teamwork**, Intellectually Curious, Continuous Learner, **Leadership & Communication Skills**, Time Management, Optimist, Motivated, Analytical Thinker, **Data-Driven Problem Solver**

Projects

Data Science and Machine Learning   *2017-2021*
- *F1 Data Analysis*: Analyzed previous Turkish GP race lap times, and visualizing race pace trends across different years
- Completed an 8-month Undergraduate Thesis ML project on Detection of Sexual Predatory Behavior in Chat-rooms
- Provided **statistical consulting** services for the *Statistical Society of Canada* and improved their conference schedule
- Performed Unsupervised **Feature Selection, Outlier Detection**, Time Series Analysis as professional development

Deep Learning   *2019-2021*
- *F1 Side-Project*: Built a **binary classifier** using **LSTM** models that predicts if a pit-stop will happen the next race lap
- Completed the TensorFlow in Practice Specialization course; fitted **MLP, CNN, RNN & LSTM** models for various tasks
- Conducted research in *Anomaly Detection* for NLP data using **Generative Adversarial Networks**; available on GitLab
- Proposed a deep learning solution for *Text Entailment and Semantic Relatedness* **NLP** task, available on GitHub