

# Norbert Eke

DATA SCIENTIST

+1 250 878 7547 | [norberteke@protonmail.com](mailto:norberteke@protonmail.com) | [norberte.github.io](https://norberte.github.io) | [GitHub](#) | [LinkedIn](#)

## Education

### Master of Computer Science - Specialization in Data Science

CARLETON UNIVERSITY

Graduated with a GPA of 11.4/ 12. [Master's Thesis](#) in NLP, Machine Learning, and Data Mining

*Ottawa, ON, Canada*

*Sept 2018 - June 2020*

### Bachelor of Science - Honours in Computer Science, Minor in Data Science

UNIVERSITY OF BRITISH COLUMBIA - OKANAGAN

Graduated with a GPA of 3.8/ 4.0. [Honours Thesis](#) in NLP, and Machine Learning

*Kelowna, BC, Canada*

*Sept 2014 - June 2018*

## Experience

### Data Scientist

HEALTH CANADA

- Analyzing compliance verification textual data in the Data Analytics and Reporting Team (DART) within the Regulatory Operations and Enforcement Branch
- Responsible for researching and developing text processing, phrase detection, keyword extraction & text summarization algorithms for compliance verification case notes

*Ottawa, ON, Canada*

*February 2021 - present*

### Graduate Researcher and Teaching Assistant

CARLETON UNIVERSITY

- Hosted office hours & graded assignments for the **Introduction to Databases** and **Distributed Computing** courses
- Conducted research on [Software Developer Expertise Learning](#) in Prof. Olga Baysal's **Software Analytics** lab
- Applied the [research](#) to **recruitment and talent search** by determining candidate expertise in **data-driven** ways

*Ottawa, ON, Canada*

*Sept 2018 - April 2020*

### Data Scientist Intern

NATIONAL RESEARCH COUNCIL CANADA (NRC)

- 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

*Ottawa, ON, Canada*

*May 2019 - August 2019*

### Undergraduate Researcher

UNIVERSITY OF BRITISH COLUMBIA - OKANAGAN

- Conducted research on **feature-based opinion mining** and delivered a [presentation](#) about our findings – [GitHub](#)
- Designed automated opinion detection from reviews about product features using **Machine Learning**

*Kelowna, B.C, Canada*

*May 2017 - August 2017*

## Skills

### Interpersonal Skills

**Teamwork**, Intellectual Curiosity, **Excellent Leadership & Communication Skills**, Time Management, Motivation, Optimism, Analytical Thinking, **Data-Driven Problem Solving**

### Programming

**PYTHON**, R, SQL, JAVA, **Object-Oriented Programming**, Data Structures, Algorithmic Thinking

### Software Engineering

Experienced with **Client-Server Applications** ([projects](#)), Algorithm Implementation, CI & CD, **Git**, Deployment, Unit Testing, **Agile Development**, Writing Reports & Documentation

### Cloud & Deployment

AWS, **Azure**, BigQuery, Flask, Docker, Experience with **ETL** pipelines

### ML Platforms and Libraries

**TensorFlow**, Keras, **SciKit-Learn**, Imbalanced-Learn, **tidyverse**, StatsModels, PyCaret, SpaCy, Seaborn, **Numpy**, **Pandas**, **Matplotlib**, **SciPy**, SQLAlchemy, Scikit-Optimize, **Gensim**, **NLTK**

### Machine Learning

**Supervised & Unsupervised**: Traditional ML (Random Forest, k-NN, SVM), **Outlier Detection**, **Time-series Forecasting**, Hyper-param. Optimization, **Regression**, **Classification**, **Clustering**

### Deep Learning

Neural Nets for **NLP** and **Computer Vision** Applications using **CNN**, **RNN**, **LSTM**, **GRU**, **GAN**

### Statistics

Descriptive Statistics, **Probability Theory**, Bayesian Statistics, **Time Series Analysis**, **EDA**, **Statistical Modeling & Inference**, Dimensionality Reduction (**t-SNE**, **PCA**), Hypothesis Testing

### Data Analysis

Extensive experience with **Data Cleansing**, **Wrangling**, **Visualization** (Tableau, Ggplot), A/B testing, **Predictive Modeling**, **Data Mining**, Handling **Unstructured Data**, **Algorithm Design**

### Databases

Extensive experience with **SQL**, Architecture Design, **MySQL**, PostgreSQL, NoSQL

## Projects

### DATA SCIENCE AND MACHINE LEARNING

*2017-2020*

- Completed an 8-month [Undergraduate Thesis](#) project on [Detection of Sexual Predatory Behavior in Chat-rooms](#), improving my **Data Analytics**, **Statistical Machine Learning**, and **Natural Language Processing** experience
- Provided **statistical consulting** 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

*2019-2020*

- Completed [TensorFlow in Practice Specialization](#) course; fitted **MLP**, **CNN**, **RNN & LSTM** models for various tasks
- Conducted research in Anomaly Detection using **Generative Adversarial Networks**; available on [GitLab](#)
- Trained models for *Text Entailment and Semantic Relatedness* **NLP** task, available on [GitHub](#)