

# NORBERT EKE

Data Scientist

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

## EDUCATION

**Carleton University**, Ottawa, ON, Canada 2018 - 2020

*M.Sc Computer Science (Data Science Specialization)*

- Graduated with GPA 11.4/ 12. Master's Thesis in NLP, Machine Learning and Data Mining.

**University of British Columbia - Okanagan**, Kelowna, BC, Canada 2014 - 2018

*B.Sc (Honours) Computer Science (Minor in Data Science)*

- Graduated with 90% in major average. Honours Thesis in Machine Learning, and Data Science.

## WORK EXPERIENCE

**Carleton University**, Ottawa, ON, Canada Sep 2018 - Apr 2020

*Graduate Researcher and Teaching Assistant*

- I performed research towards my Master's thesis titled *Cross-Platform Software Developer Expertise Learning*.

**National Research Council (NRC) Canada**, Ottawa, ON, Canada May 2019 - Aug 2019

*Data Scientist Intern*

- Completed a 4-month contract for a government client. Delivered internal report to client, presented findings, shipped source code, analysis results and demo. The project was heavy on data collection, NLP, and Bayesian statistics.

**University of British Columbia - Okanagan**, Kelowna, BC, Canada May 2017 - Aug 2017

*Undergraduate Researcher*

- Researched word embeddings applied in feature-based opinion mining. Presented findings at UBC's Research Symposium.

**University of British Columbia - Okanagan**, Kelowna, BC, Canada Sep 2015 - Apr 2018

*Undergraduate Teaching Assistant*

- Helped students apply concepts taught in lectures via hands-on programming labs and marked their assignments, midterms.

## SKILLS

<i>Programming</i>	PYTHON, R, JAVA, SQL, JAVASCRIPT, HTML, CSS
<i>Machine Learning Tools</i>	TensorFlow, Keras, SciKit-Learn.
<i>Python Libraries</i>	Experience with Pandas, Numpy, Matplotlib, Gensim, NLTK, SpaCy, Scikit-Optimize, pyLDAviz.
<i>Machine Learning</i>	Decision Tree based ML (Random Forest, Bagging, Boosting), Traditional ML (Clustering, Classification, Linear & Logistic Regression), Deep Learning/Neural Networks, Bayesian methods, Topic Modeling, Sentiment Analysis, Image Analysis, Word Embeddings.
<i>Statistics</i>	Statistical Modeling, Inference, Probability Theory, Dimensionality Reduction (t-SNE, PCA), Time Series Forecasting, Hypothesis Testing, Bayesian Change-Point Analysis.
<i>Databases</i>	MySQL, BigQuery, PostgreSQL.
<i>Data Analysis</i>	Extensive hands-on experience with Data Cleaning, Wrangling, Visualization (D3, ggplot2, Tableau), Statistical Modeling, Data Mining, EDA, Database and Algorithm Design.
<i>Soft Skills</i>	Teamwork, Leadership, Communication, Time Management, Work Ethic, Critical Thinking.

## ACADEMIC AND PUBLIC PROJECTS

**Online Courses: TensorFlow in Practice Specialization**, Coursera course May 2020

**Deep Learning: Text Entailment and Semantic Relatedness**, GitHub repository March 2019

**Machine Learning Research: Anomaly Detection using Generative Adversarial Networks**, GitLab repository Fall 2018

**Data Science Research: Identification of Sexual Predatory Behavior in Chat-rooms**, Undergraduate Thesis Winter 2018

**Statistical Consulting: Improved the conference schedule for Statistical Society of Canada**, GitHub repository Fall 2017

**Data Science Research: Word Embeddings Applied in Opinion Mining**, GitHub repository Summer 2017