Colleen Smith

Data Scientist

colleen.smith167@gmail.com
613-408-5626
ca.linkedin.com/in/colleendsmith/

Objective

I'm a skilled data scientist who has worked in the field for 5 years. My experience includes working closely with the data and applying domain knowledge to determine the best models to use for the machine learning problem at hand. I'm also skilled software developer which allows me to work well in a mixed team of data scientists and engineers. My ideal position is one where I can use my skills in data science to directly improve people's lives, such as in the domains of healthcare or environmental science.

Skills

- Statistics
- Anomaly Detection (One-class SVM, KNN, Autoencoder)
- Classification/Regression (SVM, GLM, Random Forests, KNN, NN)
- Data Mining, Cleaning, Exploration, Visualization
- Feature Engineering/Feature Selection
- NLP
- Python (Pandas, Numpy, Scipy, scikit-learn, plotly, matplotlib, keras)
- C/C++
- R (dplyr, ggplot)
- SQL/MongoDB
- Technical writing

Measurable Results

- Developed a machine learning product to detect anomalies in mouse movements made by users of a browser app. This system achieved promising initial results and helped attract investors.
- Worked closely with a client to develop a custom product that detects anomalies in mobile device usage data. This contributed to the signing of proof-of-concept contracts.
- Developed a brain-computer interface by applying machine learning models to brain data. Experiments proved this system could increase communication rates for individuals with paralysis.

Career History

Data Scientist

Zighra, Nov 2014 – Present

- Lead the development of machine learning algorithms for biometric-based security.
- Use clustering and linear models to detect outliers in data.
- Apply feature engineering and feature selection to methods to data.
- Develop Python libraries for production environments.
- Use Plotly and Matplotlib for visualization of biometric data
- Present data analysis results to business team and customers.
- Conduct experiments to collect biometric data

Researcher/Analyst
Personal Neuro Devices, May 2013 – Oct 2014

- Direct R&D of EEG algorithms to measure and predict cognitive states.
- Use statistical techniques to classify EEG data.
- Design and conduct data collection experiments.

Education

Masters in Biomedical Engineering
University of Toronto
GPA 4.0/4.0

- Use machine learning to develop brain-computer interfaces
- Use natural language processing (NLP) to create model of lexical memory

BSc in Mathematics and Engineering
Queen's University
GPA 4.05/4.3

Interests

- Ottawa Trivia league
- Knitting
- Yoga
- Zero-waste and anti-consumerism movements