

PAUL OWE

EDUCATION

B.A. Computer Science

September 2016 – June 2020

York University, Toronto, ON, Canada

Relevant Courses: Data Mining, Database Systems, Design and Analysis of Algorithms, Fundamentals of Data Structures, and Software Design

WORK EXPERIENCE

Research Assistant, York University

May 2020 – Present

- Developing a proof of concept for the automatic retrieval and archiving of **open educational resources**.
- Conducting research on tagging systems, where a logical tag hierarchy is formed, and each OER object receives metadata used for search and information retrieval.

Data Scientist, PharmAssess Diagnostics Corp.

March 2020 – Present

- I lead overall data management solutions and design for a health diagnostic startup company.
- Versed in architecting and building a solution from conception -> ETL -> modelling -> to deployment of **ML inference engines on AWS** using Lambda, Sagemaker, API Gateway and a **serverless architecture**.

Analytics Intern, KPMG Data Analytics Virtual Internship

August 2019 – April 2020

- Pro-actively communicated with a business client to identify data quality issues and ways to ensure completeness.
- Prepared a detailed approach for feature engineering, transformations, modelling and reporting.
- Leveraged data visualization tools (**Tableau**) to develop dashboards and represent data in a business-friendly manner

ACADEMIC PROJECTS

Natural Language Processing

January 2020 – April 2020

- Predicted sentiment of Yelp Business Reviews with 91% accuracy by applying **supervised machine learning** models on a large review dataset – Used the bag-of-words model for text representation.
- Evaluated model performance using accuracy scores, AUC-ROC curves and misclassification costs.
- Wrote a report to discuss and justify methods for preprocessing the data and learning algorithms used for the model.

Designing Databases for an Online Couch-surfing System

September 2019 – December 2019

- Analyzed the domain, prepared a detailed specification of how information will be represented and stored in a database management system - Used **PostgreSQL** for this application.

INDEPENDENT PROJECTS

Social Network Analysis

December 2019 – January 2020

- Performed social network analysis to understand the structural properties of networks on twitter, or, graph mining. I used tweepy to test and integrate twitter's **REST and Streaming APIs** into my application. This allowed me to model twitter conversations and monitor how users interact and communicate in an aggregate manner.
- Created a dynamic map to visualize location of tweets and enhance comprehensibility of my analysis – with Folium.

Analyzing Big Data with Hadoop

February 2020 – March 2020

- Launched a fully functional Hadoop cluster using Amazon EMR, defined schema and created a table for log data stored in Amazon S3, and analyzed the data using HiveQL.

Voice Enabled Banking Assistant

January 17 – January 19, 2020

- Created a voice application for Google Assistant and Alexa to heighten a user's digital banking experience at the University of Toronto Hackathon VII. Used Voiceflow and VoiceXML to specify voice responses and dialogs.

SKILLS

Languages: Python (nlTK, scikit-learn, matplotlib, numpy, pandas), SQL || Web frameworks: React, Flask || Cloud: Amazon Web Services || Software: MS Excel, Weka, Tableau, MATLAB, RStudio || Other: Open-source contributions (Github), Data-Driven Presentations, Agile methodology, Docker, Serverless Architecture

AWARDS AND CERTIFICATES

- **Certificate:** Machine Learning Certification (Stanford University on Coursera) **August 2020**
- **Certificate:** KPMG Data Analytics Consulting Virtual Internship **April 2020**
- **Award:** 1st Place Startup Team (Bergeron Entrepreneurs in Science and Tech.) **November 2019**

VOLUNTEERING EXPERIENCE

- **Volunteer:** Service Learning at Akorlikope (Ghana) **July 2016**