SUMMARY

Data Scientist with strong development background and 15+ years of experience using quantitative methods to solve challenging problems.

SKILLS

- Python, Docker, Pandas, Scikit-Learn, Lifelines, Falcon, Flask, NumPy, TensorFlow, OpenCV
- Java, Spark, Kafka Streams, Mahout, Akka, Elasticsearch, MongoDB
- Javascript, JQuery, Bootstrap, Plotly.js
- Data Cleaning, Modeling, and Mining
- Machine Learning, Survival Analysis, NLP, Graph Algorithms
- Deep Learning, Image Classification, Object Detection, Computer Vision, Forecasting

EDUCATION

MSC. PHYSICS

Queen's University

HONOURS BSC. PHYSICS

Lakehead University

EXPERIENCE

DATA SCIENTIST - OFFICER

Wealth Management & Investment Services, U.S. Bancorp

April 2019 - Current

- Built "flight risk" application to identify investments, accounts, and clients likely to leave the bank using survival analysis
- Developed real-time investment / account balance forecasting algorithm based on deep learning
- Built incident prediction algorithm based on Random Forest and time-series ensemble forecasting to provide early warning of stability issues in business-critical applications

SOFTWARE ENGINEER

Data Science, C.H. Robinson

April 2017 - April 2019

- Built document classification application based on deep learning and image processing to identify incoming shipping documents. In production the solution examines nearly 35,000 images per day and is 84% correct in its predictions.

- Built signature detection application based on deep learning, object detection, and image processing to identify whether a packing list had been signed by the recipient. In pre-production, the solution examines several thousand documents per day and is correct more than 96% of the time when its confidence is 60% or greater.
- Designed an algorithm for automatically extracting fields such as invoice amount and payee from OCR data without training a model using similarity metrics, Named Entity Recognition (NER), and graph edit distances.

FOUNDER

Emphysic LLC

August 2015 - April 2018

- Funded by NASA to design a distributed algorithm based on online learning, image processing, and pipeline parallelism to automatically detect structural damage in aircraft. A visualization of the algorithm is available at https://youtu.be/G95QvW-zXgQ and a demonstration of the subsequent application is available at https://youtu.be/eNno2Qa_kd8 .

DATA ANALYTICS ENGINEER

Contata Solutions

April 2015 - February 2016

- Developed proof of concept approach for finding anomalies in customer rewards card program data with BIRCH clustering, outlier detection, and Spark.
- Designed a NoSQL Data Lake based on Cassandra and Elasticsearch for aggregating customer data from multiple sources and schema into a single datastore.

COMPUTATIONAL PHYSICS PROGRAMMER

Computational Physics, Canadian Nuclear Laboratories

September 2013 - April 2015

- Built an extensible application for automatic analysis of nuclear reactor simulation data based on Natural Language Processing.

ENGINEER / SCIENTIST

NDE Division, TRI/Austin, Inc.

October 1999 - September 2013

- Patented a statistical approach to assessing the structural integrity of aircraft and helicopters in real time during flight, used to predict future health and recommend a course of action to mitigate. First successful test flight in 2006 on an F-15/E and the basis of more than $1 million in revenue for the company.

MAGNETICS ENGINEER

Pipetronix Ltd.

August 1998 - October 1999

- Developed a model of the behavior of the magnetic field around an in-line inspection tool as it moved through the pipeline, used to determine the bias in sensor readings as a function of velocity.

LINKS

- LinkedIn
- GitHub
- GitLab