ALEX MINNAAR

minnaaralex@gmail.com / alexminnaar.com / github.com/alexminnaar

EDUCATION

University College London *November 2012*
MSc in Computational Statistics & Machine Learning
London, United Kingdom

Queen's University *June 2011*
BSc in Mathematics & Engineering - Control and Computing stream
Kingston, Ontario, Canada

Upper Canada College *June 2007*
Ontario Diploma & International Baccalaureate Diploma
Toronto, Ontario, Canada

EXPERIENCE

VerticalScope, Inc March 2014 - Present
Software Engineer *Toronto, Ontario, Canada*

- Part of the Data Science team. Building large-scale, real-time natural language processing applications with unstructured social media data.
- Implementing NLP/machine learning concepts such as named entity recognition/disambiguation, topic modeling, sentiment analysis, and recommender systems.
- Using the Scala programming language with technologies such as Spark, Kafka, Elasticsearch and Cassandra.

Tucows, Inc February 2013 - March 2014
Data Scientist *Toronto, Ontario, Canada*

- Implemented statistical analysis to gain insight on customer churn.
- Developed several machine learning applications involving classification, clustering, collaborative filtering and sentiment analysis.
- Setup company's first Hadoop cluster using Amazon ec2 and Cloudera.

Advanced Algos, Inc October 2012 - December 2012
Data Analyst Intern *Toronto, Ontario, Canada*

- Performed data analysis tasks on CME and EBS data sets from Spot and FX markets.
- Wrote C++ programs to search for specific occurrences in large data sets.

RESEARCH AND PROJECT EXPERIENCE

MSc Thesis University College London

New Methods for Identifying Small Mean-reverting Portfolios

· Studying the portfolio selection problem of extracting optimally mean-reverting portfolios from multivariate time-series datasets while constraining the number of assets in the portfolio. Included comparison of approximation techniques such as greedy search, semi-definite relaxation, and genetic algorithms. Presents novel on-line learning algorithm that dynamically adjusts portfolio weights to achieve optimal mean-reversion.

Undergraduate Thesis Queen's University

Stochastic Systems with Data Rate Constraints

· A dynamical system was studied in which two noisy sensors are connected to a controller through a noiseless finite capacity channel. A stability result was presented in the case where the eigenvalues of the system matrix are real and distinct. A policy that achieves stochastic stability is also presented and is shown to asymptotically achieve the minimum average rate for the communication channel.

INVITED TALKS

Scalable Online Learning of Topic Models with Spark 4/25/2015

Text By the Bay NLP Conference

· This talk deals with the problem of how to learn topic models from large text corpora that are constantly growing such as with online forums. As documents stream into your corpus it is much more efficient to update your already learned topic model rather than batch processing your entire corpus. Apache Spark can be used to perform the sequential updates in a distributed fashion. The talk will also include a discussion on how to use your learned topic model to classify the documents in your corpus based on the topics they contain.

TECHNICAL STRENGTHS AND ACHIEVEMENTS

Programming Languages	C++, Java/Scala, Python, Javascript/JQuery.
Statistical Software	Matlab, R.
Databases	MySQL, Elasticsearch, Cassandra.
Hadoop & Big Data	MapReduce programming, Hive, Mahout, Spark.
Awards	Ontario Scholar. Upper Canada College Prefect. Euclid Math Contest Certificate of Distinction.
Certificates	Hadoop Fundamentals Certificate, *BigDataUniversity.com.* C++ for Financial Engineering, *Baruch College.*
Personal Projects	alexminnaar.com (blog), github.com/alexminnaar (open source projects).
Interests	Brazilian Jiu-Jitsu, CrossFit, Music, Movies.