

# Liam Bui

📍 Vancouver, BC

☎ (604) 352-1568

✉ [liam\\_bui@outlook.com](mailto:liam_bui@outlook.com)

🌐 <https://liambll.github.io/blog/>

## TECHNICAL SKILLS

INFORMS Certified Analytics Professional with experience in Operations Research, Statistical Analysis, Machine Learning, and Big Data.

- **Big Data:** Hadoop, Spark, Hive, Cassandra, MongoDB
- **Modeling:** Scikit-learn, Tensorflow
- **Programming:** Python, Scala, Java, R, SAS, JavaScript
- **Visualization:** Qlikview, Tableau, D3.js

## PROFESSIONAL EXPERIENCE

**Big Data Engineer (Coop) – PHEMI, Vancouver, BC** May 2017 – Present

- Develop big data processing and analytics prototypes using Spark/Scala, Hive, Zeppelin
- Implement natural language processing pipeline with Apache cTAKES (clinical Text Analysis Knowledge Extraction System) to extract clinical information from unstructured medical text
- Develop medical image analysis demo using Python machine learning libraries (Scikit-learn, Tensorflow)

**Business Analytics Associate - DBS Bank, Singapore** Jul 2015 – Jul 2016

- Provided insights on Card & Unsecured Lending 's sales and marketing by extracting data with SQL and conducting statistical analysis and data mining in SAS
- Doubled customer response rate in digital campaigns by liaising with data science team to deploy predictive analytics models (Recommender System, Location Analytics) for targeted marketing
- Improved customer engagement with email marketing (1.3x lift in email click rate) through experiment design and hypothesis testing on different content factors
- Managed vendors for the monthly delivery of Qlikview dashboards for marketing metrics tracking

**Data Analyst - DBS Bank, Singapore** Jul 2012 – Jun 2015

- Developed Java analysis reports and Qlikview dashboards to provide technology and operations teams with insight on process improvement and risk control
- Reduced operating costs by 10% through process reengineering and queuing simulation to evaluate process improvement opportunities
- Led user engagement and managed modelling team to model business processes, data and system architectures using industry standard notations, leading to consistency and standardization

**Business Intelligence Developer (Intern) - Hutcab Consulting, Singapore** May 2011 – Aug 2011

- Implemented inventory management and demand forecasting dashboards in a Decision Support System with Java Servlet.
- Reduced data processing time by automating Data Extract, Transform and Load (ETL) tasks with Microsoft SQL Server Integration Services (SSIS)

## EDUCATION

**MSc. Computer Science - GPA 4.12/4.33** Sep 2016 – Present

Simon Fraser University, Burnaby, BC, Canada

- *Big Data Systems, Big Data Analytics, Machine Learning, Deep Learning, Data Mining, Regression Analysis, Visual Analytics*

**BEng. Computer Science, BSc. Business Analytics** Aug 2008 – Jun 2012

Nanyang Technological University, Singapore

# Liam Bui

📍 Vancouver, BC

☎ (604) 352-1568

✉ [liam\\_bui@outlook.com](mailto:liam_bui@outlook.com)

🌐 <https://liambll.github.io/blog/>

## ACADEMIC PROJECT

### Topic Model and Network Analysis on research publications

Jan 2017 – Apr 2017

- Developed web scraping function to retrieval publications from major machine learning journals & conferences
- Analyzed topic models (Latent Dirichlet Allocation) to explore underlying topics and performed graph analytics to understand and visualize relationship among publication contents
- Technologies: Scala, Spark ML, Spark GraphX, MongoDB, Gephi

### Question-Answer Style Chatbot

Jan 2017 – Apr 2017

- Developed a chatbot using Sequence to Sequence Learning (LSTM) for language models, together with several improvement techniques (Beam Search, Attention Mechanism)
- Developed a web-based user interface for user-chatbot conversation
- Technologies: Python, NLTK, Tensorflow, Flask

### Longitudinal study: Effect of tobacco use on mortality

Jan 2017 – Apr 2017

- Analyzed National Longitudinal Mortality dataset to understand the effect of tobacco use on mortality using mixed model logistic regression and survival analysis (Kaplan-Meier estimator)
- Technologies: R, ggplot2, glm, lme4, survival

### All State Claim, Kaggle Data Science Competition

Sep 2016 – Dec 2016

- Predicted car accident's insurance claim using Linear Regression, Neural Network and Boosted Tree
- Ranked top 20% in the Kaggle competition
- Technologies: Python, scikit-learn, Keras, XGBoost

### Online Restaurant Recommender System

Sep 2016 – Dec 2016

- Implemented Collaborative Filtering Recommender System (Matrix Factorization, Item-Similarity) based on restaurant detail and user rating datasets from Yelp
- Developed a web-based Proof of Concept to demonstrate and visualize the recommender system's result
- Technologies: Scala, Spark, MLlib, Cassandra, Hadoop Cloudera Distribution, Google Map API, D3.js

### Facial Expression Recognition in Real-Time

Sep 2016 – Dec 2016

- Implemented and Benchmarked different models (Principle Component Analysis, Linear Discrimination Analysis, Convolutional Neural Network) to extract features from facial expression images and predict emotion in real-time
- Technologies: Python, OpenCV, Keras, Amazon AWS

### Sentiment Analysis on Movie Comments

Sep 2016 – Dec 2016

- Performed natural language processing and text mining on 6,000 movie comments using Convolutional Neural Network in Matlab to analyze users' attitudes towards movies
- Technologies: Matlab, Matconvnet, Word2Vec

## CERTIFICATIONS & TRAINING

- Certified Analytics Professional - Institute for Operations Research and the Management Sciences (2017)
- Applied Analytics Using SAS Enterprise Miner - SAS (2015)
- ANOVA, Regression, and Logistic Regression - SAS (2015)