# Junn Hei Jonathan Cho

University of Waterloo | Management Engineering

linkedin.com/in/jon88cho | github.com/jon88cho | jon88cho.github.io

## Summary of Qualifications

- Implemented accurate Machine Learning applications using Python in projects done independently, as a Research Assistant and at Home Depot
- Competency with Data Science and Statistical principles acquired through extensive dataset analyses and trend recognition within personal projects

**Languages:** Python, Java, SQL, Javascript, HTML, CSS, Bash, Microsoft Office Suite, Matlab

**Tools:** Google Cloud Platform, SAP BW & HANA, Tableau, Git, Jira, MongoDB

## Relevant Experience

**Developer, Data Analytics | The Home Depot Canada** **Jan 2018 – Apr 2018**

- Developed an interactive dashboard that predicts problems with 90% accuracy by employing Machine Learning principles in Python
- Created SQL queries and contributed logic to Java backend to support sales analysis and order journey efficiency streamlining
- Rectified $500,000 of sales discrepancies by correcting SQL queries for 15 reports by validating against an SAP BW database and participating in Scrum meetings

**Research Assistant | University of Waterloo** **Sept 2017 – Present**

- Applying cutting-edge Machine Learning techniques (Adversarial Learning, Multi-Task Learning) in TensorFlow for product suggestion system with $60,000 of funding from Amazon
- Analyzed and tested advanced models from Natural Language Processing literature like RNNs, Encoder-Decoder architectures, Variational techniques and Word Vectors

## Select Projects

**Help Desk Ticket Analysis Dashboard** | *Python, Tableau*

*The Home Depot Canada*

- Cleaned large amounts of unstructured linguistic data using NLTK and Pandas
- Explored semantic data in tickets using DBSCAN and Word Vectors from SciKit-Learn and Gensim
- Devised a model using LSTMs and a feed-forward architecture in Keras that classifies help desk tickets based on semantic content with 90% accuracy
- Created interactive visualizations of geographic and time variables in Tableau

**Chat-Bot** | *Python* | Link

- Implemented Variational Recurrent Encoder-Decoder model from literature with Keras
- Cleaned Cornell Movie-Dialog Corpus into sequences of Word Vectors using Pandas and Gensim
- Train and evaluated the model's ability to generate human-like responses

**Basketball Data Exploration Project** | *Python* | Link

- Mined Basketball Data using Pandas collected from various websites using Beautiful Soup
- Explored data using Matplotlib and Seaborn with Box plots and Multivariate Distributions
- Performed statistical analyses like Chi-Squared Testing and confidence intervals to find correlations between NBA players' defensive statistics and perception of defensive skill