MOYAN MEI

Cell: 604-725-9859 ◇ **Email**: mmy12580@gmail.com ◇ **Website**: moyan.ml
Fashion District, Toronto, Ontario, Canada

EDUCATION

Simon Fraser University *2014.9 - 2016.8*
Master of Science in Statistics Burnaby, Canada

- Graduate Fellowship, GPA: 3.76/4.0

Dalhousie University *2011.5 - 2014.5*
Bachelor of Science in Statistics (Honors) Halifax, Canada

- **Highest** GPA: 3.92/4.0 among (25+) major courses
- First Class Honors, cGPA: 3.82/4.0
- President's Entrance Scholarship, cGPA: 92/100

CORE QUALIFICATIONS

Language & Tool	Python, R, SQL, Spark, Matlab, MongoDB, Bash, LaTeX Google Cloud Service, AWS S2/EC2, Docker
Deep Learning Framework	Pytorch, Tensorflow, Keras, PaddlePaddle, Mxnet, Theano
NLP Library	Allennlp, Flairnlp, Fastnlp, Hanlp, Spacy, NLTK
In-depth Knowledge	Deep Neural Networks, Machine Learning methods, Optimization Error Analysis, Statistical Inference, Probability, Linear Algebra

EXPERIENCE

WGames Inc *2018.1 - Present*
Machine Learning Scientist Toronto, Canada

- Leveraged hybrid latent representation learning models e.g. lightFM and Variational AutoEncoder with WARP loss to recommend high accuracy $\approx 90\%$ daily games to half-million customers
- Designed topic classification, sentiment classification and smart-reply models (> 5 million reviews) with distilled BERT model to improve user experience
- Created a scalable and flexible machine learning pipeline, e.g imbalanced supervised learning and unsupervised learning, for customer segmentation to identify their unmet needs
- Recommended business insights from retention analysis, engagement analysis and non-fatal error analysis which leads to **Top 3 most popular in Google Play** under the same category

Leafy AI *2017.12 - 2019.7*
NLP Scientist Beijing, China

- Developed a Chinese multi-task learning Python toolkit for customers' daily efficiency, e.g., word segmentation, named entity recognition, event extraction, and topic classification
- Boosted in-app search experience of customers by training semantic similar embeddings (e.g words, sub-words, and sentence based) with scalable and memory-efficient FAISS
- Built a transferable and configurable personalized Question and Answer (Q&A) system for customers to solve their needs from their knowledge database
- Implemented knowledge distillation and compression on deep learning NLP models to obtain low-memory and mobile-friendly offline models in ONNX format

Istuary Innovation Group *2016.9 - 2017.10*
Data Scientist Vancouver, Canada

- Designed a 1:1 facial verification, 1:N facial recognition, and face alignment deep learning prototypes embedded in smart camera
- Proposed a two-stage facial verification method, which improves the state-of-the-art model by 6% - 30% on different case scenarios
- Maintained a fast, high-quality, and large-scale image data pre-processing framework, including image cropping, resizing, clustering, and augmentation
- Established an automatic summarization API which produces a condensed representation of its inputs for the Chinese news APP by adapting KL divergence, TextRank, and Recurrent Neural Networks
- Increased customer stickiness to our news APP by recommending semantic similar news with fastText

Center for Operations Excellence, UBC (Co-op) *2015.5 - 2015.9*
Technical Analyst/Statistician Vancouver, Canada

- Applied sentiment analysis with Naïve Bayes, Random Forest, and Hierarchical Clustering, with TF-IDF on over 8 Million Tweets
- Analyzed negative sentiments and provided useful insights (i.e., flight delay, customer service, uncomfortable seats, and poor food supplies) to Boeing company by applying Latent Dirichlet Allocation
- Deployed a tweets processing pipeline including collecting, pre-processing, classifying sentiments and extracting topics embedded in a dashboard

COMPETITION

Statistical Society of Canada Conference Competition *2014.1 - 2014.5*
Uken Company Toronto, Canada

- Applied exploratory analysis with visualization on 300K users to obtain interpretable features about revenue among predictors, e.g gender, platform, and in-game items, etc.
- Built an ensemble high accuracy $\approx 94\%$ classification model from logistic regression, linear discriminant analysis, and support vector machine for the retention of the game users
- Constructed additive regression models, e.g. generalize additive model and multiple linear regression to predict overall revenues with accuracy $\approx 87.5\%$
- **3rd place** winner of the case study competition

ACHIEVEMENTS

Graduate Fellowship from Simon Fraser University *2016.1*
Third place winner of SSC Conference Competition *2014.5*
Six times Dean's List at Dalhousie University *2011.5 - 2014.5*