Yong Zhang, Ph.D.

Homepage:	nickyongzhang.github.io
Email:	nickzylove@gmail.com
Phone:	+1 2066982844
Address:	1117 N 199th St., Shoreline, WA, 98133

—Education—

08/2013-07/2017 Ph.D., School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore; GPA – 4.92/5.0

Thesis title: *Semantic Representation Learning for Natural Language Understanding*

09/2009-06/2013 B.Eng., School of Electrical Engineering and Automation, Zhejiang University, China; GPA – 3.88/4.0; *First-class Honor graduate*

—Experience—

04/2019-Current Research Scientist (Amazon)

Owned private brands recommendation at detail pages project. One product of the project contributes 10's of millions of dollars to Amazon.

- Proposed a personalized recommendation model based on temporal convolutional neural networks which capture customers' sequential behavior patterns. The model is developed using Pytorch. The paper of the work has been submitted to AMLC 2020.
- Participated in building an AWS-based machine learning system that fully automates all aspects of the model from data extraction, model training to recommendation serving. The system serves real-time recommendations to customers at detail pages taking into consideration both context page asin and customer id. Tools used include SageMaker, Amazon Step Functions, Amazon ElastiCache, Docker, etc.
- Delivered research talk at *Global Softlines Science Summit - 2020* and *Prime Machine Learning Talk Series*.

Owned ads performance models of Amazon internal advertisers. The project enables Amazon private brands products to exit common sponsored product auction so they do not directly compete against external advertisers.

- Developed Gradient Boosting based models to estimate CTR and CVR values of private brands products shown in ads slots. Deep learning models, including DeepFM, xDeepFM, AFM, AutoInt, etc., have also been explored and experimented.
- Wrote ETL package using AWS Glue. Scala is the necessary skill.
- Expanded the model to international markets.

Main contributor of private brands recommendation at search pages project. The project contributes over 100 million dollars to Amazon private brands.

- Productionize a neural information retrieval model named Deep Structured Semantic Model for retrieving relevant products for search queries. The production code is 10 times faster than prototype code. Tensor flow is the main framework used.
- Expanded the model to international markets, including UK/DE/FR/ES/IT/JP/IN/CA/MX. Wrote a utility package to process non-English languages. SpaCy is utilized.

- Participated in building a fully automated machine learning pipeline from model development to production.

07/2018-03/2019 Data Scientist (Shopee)

Developed Unique User Identification System - a graph-based identity management system that link accounts belonging to the same person, aimed at fraud detection, credit scoring, marketing, etc.

- Built graph infrastructure and graph DB from scratch. Janusgraph, Apache Tinkerpop, Pyspark and Hadoop are basic tools utilized.
- Explored methods for large-scale graph data ingestion and fast graph query. Skills in Gremlin and Elasticsearch are necessary.
- Designed graph embedding algorithms to effectively represent users with numeric vectors so that machine learning models can be employed for downstream tasks, like fraud detection and credit scoring.
- Designed and experiment efficient searching method for fuzzy matching of similar entities. Knowledge in the area of hashing and Redis data storage are used.

Was in charge of Categorization Recommendation Project - an automatic system making category suggestions to sellers when listing products. The model is deep learning-based and serves up to 2m calls per day in TW market alone.

- Designed machine learning models to do category suggestion based on product title. Cutting-in edge algorithms, like Convolutional Neural Networks, Recurrent Neural Networks, Fasttext, Transformers, and Bert are explored and experimented.
- Developed Web-Api service and deploy developed models to live environment. The live service has to endure tens of millions of calls every day from sellers of seven countries across Southeast Asia and Taiwan. Some basic tools used include Flask, Gunicorn and Logstash.
- Dealt with all kinds of issues arisen from the project related with model performance, API calling, and business requirements.
- Coordinated with stake-holders on the live service deployment plan of seven countries.

07/2017-06/2018 Data Scientist (GroupM)

Developed mPID in-house solution - the foundational cross-device and cross-platform consumer identity management system that ties individual's identities across devices into a consumer ID for audience re targeting across all media.

- Performed ETL and effectively integrated data from various sources including self-owned data, agency data, client data, and third-party data. Hive is mainly utilized.
- Designed and selected features and built pair scoring engine to identify device pairs belonging to the same individual.

Worked on Consumer Insights Project - the bridge connecting research and marketing whose purpose is to understand and interpret non-obvious mindsets, desires, and motivates that trigger consumers' attitudes and actions towards a brand which aims to enhance marketing effectiveness and increase sales for mutual benefit.

- Explored and integrated data from various sources and visualized the data for clear and in-depth understanding.
- Developed web-based interactive app. R shiny and python plotly are used.
- Designed appropriate attribution and clustering model for consumer journey and cluster analysis.

08/2013-07/2017 Ph.D. researcher (Nanyang Technological University)

Designed and developed several innovative machine learning algorithms with specific applications in natural language understanding tasks, including document summarization, sentence classification, sentiment analysis, etc. Published three journals, four conference papers, and two book chapters.

- Developed an algorithm named Multi-view Convolutional Neural Network and applied to multi-document summarization. The work is published in IEEE Transactions on Cybernetics.

- Proposed an innovative pooling strategy for Convolutional Neural Network based on attention mechanism and applied to sentence classification. The work is published in Information Sciences.
- Designed an active learning model based on extreme learning machine and meta-cognitive learning. The work is published in Neurocomputing.

07/2016-12/2016 Intern (GroupM)

Participated in the Audience Demography Estimation project which uses machine learning models to predict audience demography based on users' browsing logs for targeted advertising.

- Built web crawler to parse webpages and scrape information.
- Designed feature representation models.
- Employed various advanced machine learning models to predict audience demography attributes, (i.e., gender, age, house income).
- Performed ETL for large volumes of online user-level data for data cleansing and preparation.

08/2015-12/2016 Teaching Assistant (Nanyang Technological University)

Performed grading, supervised lab sessions, evaluated final presentations for courses in EEE undergraduate programs in Electric Field, Magnetic Field, Transformer, and Electrical Motor.

—Skills—

Adept: Tensorflow, Pytorch, Python, R, SQL, Pyspark

Basic: C, C++, HTML, JavaScript, Scala

—Honors—

Nanyang Research Scholarship, 2013-2017
Zhejiang University Excellent Student Awards, 2010-2012
First-class Scholarship for Outstanding Student (top 5%), 2011–2012
Wangguosong Scholarship (the most prestigious honor of the school), 2012
Sichuan Province Excellent Student Leader, 2009

—Language—

Chinese: Native || English: Professional working proficiency

—Certificates—

Passed Level 1 of the CFA exams, Google Adwords Certificates, Google Analytics Certificates

—Interest—

Reading, Travelling, Playing basketball, Swimming, Fitness

—Selected Publications—

[J1] **Yong ZHANG**, Meng Joo ER, Rui ZHAO and Mahardhika PRATAMA, "Multi-view Convolutional Neural Networks for Multi-document Extractive Summarization", *IEEE transactions on cybernetics* 47.10 (2017): 3230-3242.

[J2] Meng Joo ER, **Yong ZHANG***, Ning WANG and Mahardhika PRATAMA. "Attention Pooling-based Convolutional Neural Network for Sentence Modelling", *Information Sciences*, 373(2016): 388-403. (*corresponding author)

[J3] **Yong ZHANG** and Meng Joo ER. "Sequential active learning using meta-cognitive extreme learning machine", *Neurocomputing*, 173(2016): 835-844.

[J4] **Yong ZHANG**, Hongming Zhou, Nganmeng Tan, Saeed Bagheri, Meng Joo Er. "Targeted Advertising Based on Browsing History", arXiv preprint arXiv:1711.04498 (2017).

[C1] **Yong ZHANG**, Meng Joo ER, Rajasekar Venkatesan, Ning WANG and Mahardhika PRATAMA, "Sentiment Classification Using Comprehensive Attention Recurrent Models", in *Proceedings of International Joint Conference on Neural Networks (IJCNN)*, Canada, 2016:1562-1569.

[C2] **Yong ZHANG**, Meng Joo ER, "An Extracted Document Summarization Framework Based on Convolutional Neural Networks", in *Proceedings of 42nd Annual Conference of IEEE Industrial Electronics Society (IECON)*, Florence, 2016: 918- 922.

[C3] **Yong ZHANG**, Meng Joo ER and Rui ZHAO, "Multi-document Extractive Summarization Using Word Embedding based Extreme Learning Machine", in *Proceedings of Computational Intelligence, IEEE Symposium Series on*, South Africa, 2015: 404-410.

[C4] **Yong ZHANG**, Meng Joo ER, Sundaram SURESH. "Meta-cognitive fuzzy extreme learning machine", in *Proceedings of Control Automation Robotics & Vision (ICARCV), 13th International Conference on. IEEE*, Singapore, 2014: 613-618.

[C5] **Yong ZHANG**, Mary MA, Simon BULL, Yvonne WANG, Howard LIN, Sandeep ATLURI, "Sequential Recommendation based on Feature-aware Temporal Convolutional Networks", submitted to *Amazon Machine Learning Conference 2020*.

[C6] Amy RUSCHAK, Simon BULL, Daniel LU, **Yong ZHANG**, Howard LIN, Mainak DATTA, Param BIDJA, Sandeep ATLURI, "A Pair of Neural-IR Models for retrieving Relevant ASINs for Search Queries", submitted to *Amazon Machine Learning Conference 2020*

[B1] Meng Joo ER, Fan LIU, Ning WANG, **Yong ZHANG**, Mahardhika PRATAMA, "User-Level Twitter Sentiment Analysis with a Hybrid Approach". in *Advances in Neural Networks – ISNN 2016*. Lecture Notes in Computer Science, vol 9719. Springer

[B2] Meng Joo ER, and **Yong ZHANG**. "Adaptive Modeling and Intelligent Control of a Sodium Nitroprusside Delivery System." in *Artificial Neural Network for Drug Design, Delivery and Disposition*. 2016. 333-354.