Yan **Yan**

Engineering Manager · Tech Lead · ML Engineer

(+1) 617-8211439 | chris.yan.y@gmail.com | scholar.google.com/citations?user=rmMWS_kAAAAJ&hl=en | www.linkedin.com/in/chrisyany/

SUMMARY

My focus areas include online advertising modeling and optimization, ads recommendation, large scale ranking models, ads delivery system, natural language processing & crowd-sourcing. I have worked in multiple companies as MLE and TLM, currently I am working at Tiktok / Bytedance as the engineering manager in charge of shop-tab ads (ttmall) recommendation backend delivery.

WORK EXPERIENCE

Tiktok / Bytedance

San Jose, USA

ENGINEERING MANGER AT SHOP ADS TEAM (10 PPL.) *Sept. 2021 - Present*

- **shop-tab recommendation delivery flow**: Spearheaded the backend delivery team to scale shop-tab recommendation ads (product listing ads), driving daily revenue from 0 to **$1.4M** globally by optimizing ad algorithms and targeting strategies. Achieved over 10% revenue take rate in SEA and US, with GMV penetration exceeding 40 % globally.
- **catalog listing ads**: Directed the development of open-loop catalog list ads in TikTok feeds, integrating products from multiple advertisers into a single engaging ad. Boosted user interaction and achieved a revenue run rate of approximately $50K USD through targeted ad placements and optimization.

Facebook / Meta

Menlo Park, USA

TECH LEAD AT FORMAT ADS RANKING TEAM *Dec. 2018 - Aug. 2021*

- **ads dynamism and ads autoflow**: Led the development of Facebook's largest auto-format product series, Autoflow, directly supported by the organizational VP, targeting a 5% revenue increase within 3 years. Oversaw the creation of backend infrastructure and innovative features like image filtering, video auto-cropping, and comments re-ranking, contributing to a 3% statistically significant revenue uplift.
- **dynamic creative ads multi-field optimization**: PoC of DCO (Engineered counter and model-based methods for text asset optimization in the Dynamic Creative Optimization (DCO) project, driving over 1% of Facebook's revenue growth. Mentored and developed two team members, leading to their successful promotions.
- **dynamic rendering of linked page post ads**: Designed and implemented machine learning models and backend infrastructure for the dynamic rendering of Linked Page Post (LPP) ads, resulting in a 0.3% increase in total Facebook app revenue.
- **dynamic rendering of ads primary text body**: Developed machine learning models and backend systems for dynamically rendering the primary text of Facebook app ads, achieving a 0.4% increase in total app revenue.

JD.com

Beijing, CHN & Mountain View, USA

SR. MANAGER AT ADS RECOMMENDATION TEAM (30 PPL.) *Mar. 2017 - Dec. 2018*

- **main app recommendation improvement**: Led the ads recommendation team in enhancing data generation, feature extraction, and model redesign by leveraging diverse data distributions and statistical patterns of app positions. These improvements resulted in a 7% increase in CTR and a 10% boost in ad revenue during H1 2018.
- **ads ranking and pricing**: Spearheaded the redesign of ranking and pricing strategies for ads recommendation, successfully decoupling GMV from ad revenue impacts. Reformulated the challenge as a Multiple Objectives Optimization (MOO) problem, significantly increasing ad revenue through optimized algorithms while preserving user experience.
- **autoML for deep learning model auto-update**: Directed the development of a self-evolving AutoML framework for JD's ads ranking models (wide & deep neural networks), enabling autonomous model updates and continuous self-improvement without human intervention.
- **ads ranking online learning based on contextual bandits**: Led a team of three to develop and productionize an online learning contextual bandits algorithm, resulting in a 10%+ increase in GMV for main feed ads recommendation in Q2-Q3 2017.
- **i2i retrieval based on session embedding**: I generated the product embeddings and the corresponding new retrieval based on user click sessions. The embeddings are incorporated into ranking models for boosting the model performance.

LinkedIn

Sunnyvale, USA

SR. MACHINE LEARNING ENGINEER *Sept. 2015 - Mar. 2017*

- **search by job**: Developed the 'search by job' feature, a next-generation product for LinkedIn Recruiter Search, responsible for over 40% of the company's total revenue. This feature decomposed new job postings into relevant skills, experience, and education, matching candidates to extracted requirements with greater accuracy.
- **talent match**: Enhanced LinkedIn's core Recruiter Search product, 'Talent Match,' increasing CTR by over 3% and saving tens of millions of dollars through backend infrastructure migration.
- **search by ideal candidate**: Led the development of 'search by ideal candidate,' LinkedIn's flagship product in 2016-2017, revolutionizing recruiter search by allowing searches based on ideal candidate templates. Oversaw product model training, feature and data organization, and model evaluation.

Yahoo! Labs

Sunnyvale, USA

Research Scientist *Dec. 2012 - Sep. 2015*

- **Yahoo! website traffic forecast**: Developed an R-streaming based traffic forecasting system for Yahoo!'s main sections, providing daily reports on page views and revenues with a mean absolute percentage error under 5%. This system played a critical role in guiding the company's quarterly financial reports.
- **Yahoo! ads bidding smart pricing**: Designed a smart-pricing regression algorithm based on second price + Δ, which significantly increased ad revenue for Yahoo! mobile search.

LANGUAGE & KNOWLEDGE

Coding Languages

- Python; Java, C++; SQL

Background Knowledge

- Deep Learning, Artificial Intelligence, Ads Recommendation, Information Retrieval, Natural Language Processing

Education

EECS Department, Northeastern University — *Boston, USA*

Ph.D of Science — *Sept. 2006 - Oct. 2012*

- Thesis: Learning from Imperfect and Related Labels

EECS Department, Tsinghua University — *Beijing, China*

Bachelor of Science — *Sept. 2002 - Jun. 2006*

PAPERS & PATENTS

Articles in Peer-Reviewed Journals

- Learning from Multiple Annotators with Varying Expertise. Yan Yan, Rómer Rosales, Glenn Fung, Ramanathan Subramanian, Jennifer Dy, Machine learning 95.3 (2014) pp. 291–327. Springer, 2014
- Modeling Multiple Annotator Expertise in the Semi-Supervised Learning Scenario. Yan Yan, Romer Rosales, Glenn Fung, Jennifer Dy, arXiv preprint arXiv:1203.3529 (2012). UAI, 2012

International Peer-Reviewed Conferences/Proceedings

- A Practical Deep Online Ranking system in E-Commerce Recommendation. Yan Yan, Zitao Liu, Meng Zhao, Wentao Guo, Weipeng P Yan, Yongjun Bao, ECML, 2018
- A Flexible Forecasting Framework for Hierarchical Time Series with Seasonal Patterns: A Case Study of Web Traffic. Zitao Liu, Yan Yan, Milos Hauskrecht, SIGIR, 2018
- From Query-By-Keyword to Query-By-Example: LinkedIn Talent Search Approach. Viet Ha-Thuc, Yan Yan, Xianren Wu, Vijay Dialani, CIKM, 2017
- Optimizing Gross Merchandise Volume via DNN-MAB Dynamic Ranking Paradigm. Yan Yan Wentao Guo, Meng Zhao, Jinghe Hu, Weipeng Yan, IJCAI-17 Workshop AI Applications in E-Commerce, 2017
- Active Learning from Multiple Knowledge Sources. Yan Yan, Rómer Rosales, Glenn Fung, Faisal Farooq, Bharat Rao, Jennifer G Dy, AISTATS, 2012
- Active Learning from Crowds. Yan Yan, Glenn M Fung, Rómer Rosales, Jennifer G Dy, Proceedings of the 28th international conference on machine learning, ICML, 2011
- Modeling Annotator Expertise: Learning When Everybody Knows A Bit of Something. Yan Yan, Rómer Rosales, Glenn Fung, Mark W Schmidt, Gerardo H Valadez, Luca Bogoni, Linda Moy, Jennifer G Dy, AISTATS, 2010
- Medical Coding Classification by Leveraging Inter-Code Relationships. Yan Yan, Glenn Fung, Jennifer G Dy, Romer Rosales Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2010
- Convex Principal Feature Selection. Mahdokht Masaeli, Yan Yan, Ying Cui, Glenn Fung, Jennifer G Dy, SIAM ICDM, 2010

Patent and Defensive Publications

- Using Log Data To Train for Automated Sourcing. Vijay Dialani, Yan Yan, Attorney Docket No.: 901992-US-NP (3080.H86US1) Customer No.: 45839, 2017
- Content Aware Dynamic Candidate Pool Retrieval and Ranking. Qi Guo, Yan Yan, Bo Hu, Attorney Docket No.: 3080.I76US1 Customer No.: 45839, 2017
- Smart Suggestions for Query Refinements. Xianren Wu, SP Kanduri, Vijay Dialani, Ye Xu, Yan Yan, Viet Thuc Ha, US Patent App. 15/188,590
- Query Building for Search by Ideal Candidates. Ye Xu, Viet Thuc Ha, Xianren Wu, SP Kanduri, V Dialani, Yan Yan, US Patent App. 15/168,811
- Ideal Candidate Search Ranking. Viet Thuc Ha, Yan Yan, Xianren Wu, SP Kanduri, Vijay Dialani, Ye Xu US Patent App. 15/168,903
- Dynamic Alteration of Weights of Ideal Candidate Search Ranking Model. Xianren Wu, Ye Xu, SP Kanduri, Vijay Dialani, Yan Yan, Viet Thuc Ha, US Patent App. 15/169,413
- Generation of Training Data for Ideal Candidate Search Ranking Model. Yan Yan, Viet Thuc Ha, Xianren Wu, SP Kanduri, Vijay Dialani, Ye Xu, US Patent App. 15/169,346
- Probabilistic Modeling for Hierarchical Time Series Forecasting. Yan Yan, Jian Yang, IPCOM000239722D, 2014
- Automatic Labeler Assignment Using A Model Built from Multi-Labeler Data. Yan Yan, Glenn Fung, Romer E Rosales, Jennifer G Dy US Patent 8,027,939, 2011