Yuyan Wang

Google Brain
1600 Amphitheatre Parkway Building 41
Mountain View, CA 94043

Email: yuyanw@google.com
Homepage: https://yuyan-wang-princeton.github.io

EDUCATION

Princeton University, Princeton, NJ *Sept. 2012 - June 2016*
- **Ph.D.** in Statistics, Department of Operations Research and Financial Engineering
- Thesis: Robust High-Dimensional Regression and Factor Models. GPA: 4.0/4.0
- Advisor: Prof Jianqing Fan

University of Science and Technology of China (USTC), Hefei, China *Sept. 2008 - July 2012*
- **B.S Honors** in Statistics, [Special Class for the Gifted Young](#) (4-year undergrad program for talented youths under 16 years old). GPA: 3.95/4.0 (94.53/100, 4.12/4.3); Rank: **1st / 188**
- Awards: Guo Moruo Scholarship (<1%, highest award for undergrad) and National Scholarship (<1%, twice)

RESEARCH INTERESTS

Topics: Machine Learning and Personalization, Consumer Modeling, Recommender Systems, Long-Term Value Optimization, Algorithmic Fairness

Methodologies: Deep Learning, Reinforcement Learning (RL), Statistical Machine Learning, High-Dimensional Statistics, Causal Inference, Field Experiment, Big Data Analytics

INDUSTRY EXPERIENCE

Google Brain, Mountain View, CA *Oct. 2019 - Present*
- **Senior Researcher** on Google Brain Recommender Research Team, **Tech lead** on consumer understanding and long-term value optimization for Google recommendation products. **Co-lead** on *"Assistive Machine Learning for Long-Term User Journeys" Moonshot* (a working group of **16** researchers and ML engineers), as part of a highly selective program at Google Brain for long-term audacious initiatives. Selected projects:
  Surrogate for Long-Term Consumer Experience in Recommender Systems
  - Developed a framework to identify sequential and temporal consumer behavior patterns that are predictive of long-term consumer experience in recommender systems, which is a sparse, noisy and long-horizon signal that is hard to optimize directly. Online large-scale field experiments on an RL-based recommendation system demonstrated significant improvements in key business metrics including consumer growth and retention, achieving 20% of the annual goal of a 10-person team.
  - **First-authored paper** [5] accepted to KDD (top CS conference) 2022. Quote from a Vice President at Google: "This is an excellent paper. I think the entire Core Experiences team (900+ employees) would benefit from reading it, and I would like to send it to everyone in the Core Experiences team."
  Understanding and Modeling Consumer Intent for Long-Term Optimization

- Developed a framework for extracting consumer intents (e.g. exploration or variety-seeking intent) on the personalized platforms, and designed RL policies to plan accordingly at longer time horizons.
- Designed a hierarchical reinforcement learning (HRL) based method for learning consumer intents at different temporal abstraction levels (e.g. user-level intent vs. session-level intent).
- **Invited talk at Recsys 2022** (top conference in Recommender Systems). Work under deployment at Google; Paper draft in preparation.

Optimizing Long-Term Consumer Experience as a Fairness in Multi-Task Learning Problem
- Studied the underexplored trade-offs between fairness and accuracy in multi-task learning; Proposed a data-driven fairness mitigation framework, *MTA-F*, and a set of metrics that captures the multi-dimensional fairness-accuracy trade-offs uniquely presented in multi-task problems. **First-authored paper** [10] accepted to KDD 2021.
- Demonstrated an intriguing insight that larger models are not always better for multi-task learning, and there exists a trade-off between minimizing task training conflicts and improving multi-task generalization, which has been long ignored. Proposed a data-adaptive framework for optimizing the trade-off. **First-authored paper** [7] accepted to the WebConf 2022.

Uber Technologies Inc., San Francisco, CA   *Sept. 2016 - Sept. 2019*
- *Senior Applied Scientist: 02/2018 - 09/2019; Applied Scientist II: 09/2016 - 01/2018*
- **Tech lead** on Uber Eats home feed ranking and recommendation; Founding member of Uber Eats Data Science team which became a team of 80+ during my time there. Selected projects that generated significant business impact and were deployed globally at Uber:

Multi-Objective Recommendation for a Three-Sided Marketplace
- Developed a personalized multi-objective optimization framework for Uber Eats restaurant recommendation, which optimizes for the three-sided marketplace consisting of consumers, restaurant partners and delivery partners.
- Online A/B experiments showed significant increases in consumer retention (+0.7%), basket value (+0.5%) and orders (+0.8%) for global markets, which translate to **$1.3M weekly gain** in revenue. My work was deployed globally at Uber Eats.
- **Patented** the work as first author.
- **Media Coverage: First-authored** tech blog was selected as top 10 machine learning articles of the month (0.7%) by an independent publisher. Won "Most Impressive Business Impact" award by Uber.

Holistic Optimization with Heterogeneous & Hierarchical Contents
- Proposed and developed *HRank*, a holistic recommendation framework for personalized optimal homepage layout, combining machine learning and probabilistic modeling for consumers' browsing behavior on heterogeneous and hierarchical contents.
- *HRank* was deployed globally, which brought a significant increase in consumer conversion rate (+1.5%), amounting to **$1.1M weekly gain** in revenue.
- Presented the work at two external Meetups. My work was featured in Uber's company-wide Machine Learning Orientation video which was circulated among 500+ ML engineers and applied scientists.

Microsoft Research, Redmond, WA   *June 2015 - Aug. 2015*
- *Research Intern* at Internet Services & Research Center

Morgan Stanley, New York City, NY   *June 2014 - Aug. 2014*
- Strategies & Modeling *Summer Associate*

Chinese Academy of Sciences (CAS), Beijing, China   *Feb. 2012 - June 2012*

- ***Research Intern*** at Academy of Mathematics and Systems Science

University of California, Los Angeles (UCLA), Los Angeles, China   *July. 2011 - Sept. 2011*
- ***Cross-disciplinary Scholars*** *in Science and Technology (CSST) program,* winner of CSST Award (6/90)

MANUSRIPTS UNDER REVIEW

[1] **Wang, Y.**, Tao L., Zhang X.. "Recommending for a Three-Sided Food Delivery Marketplace: A Multi-Objective Hierarchical Approach."
Major Revision *(1st round review) at Marketing Science.*
Best Paper Award Nominee *at [CIST 2022](#).*

JOURNAL PUBLICATIONS

[2] Li, Q., Cheng, G., Fan, J., **Wang, Y.** (2018). "Embracing the Blessing of Dimensionality in Factor Models." Journal of the American Statistical Association, 113.521 (2018): 380-389. (**JASA**).
Top-tier journal in Statistics.

[3] Fan, J., Li, Q., **Wang, Y. (Alphabetical order)** (2017). "Estimation of High-Dimensional Mean Regression in Absence of Symmetry and Light-tail Assumptions." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 79.1 (2017): 247-265. (**JRSS-B**).
Top-tier journal in Statistics. 2nd-highest cited paper *among all papers from the same issue. Authors are alphabetically ordered.*

[4] Lin, N., Jing, R., **Wang, Y.**, Yonekura E., Fan, J., Xue, L. (2017). "A statistical investigation of the dependence of tropical cyclone intensity change on the surrounding environment." **Monthly Weather Review**, 145 (7), 2813-2831.
Top-tier journal in Atmospheric Science.
Collaboration with the Civil & Environmental Engineering Department at Princeton University.

TOP-TIER CONFERENCE PUBLICATIONS

[5] **Wang, Y.**, Sharma, M., Badam, S., Xu, C., Sun, Q., Richardson, L., Chung, L., Chi, E.H., Chen, M.. "Surrogate for Long-Term User Experience in Recommender Systems." Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery & Data Mining **(KDD 2022)**.
Top conference in Computer Science. *Full paper with oral presentation; Acceptance rate: 15.0%.*
Work was underlined by [an invited talk at KDD 2022](#).
*Short version accepted to BayLearn 2022 (**top** ML conference in the Bay Area).*

[6] **Wang, Y.**, Tao L., Zhang X.. "Recommending for a Multi-Sided Marketplace with Heterogeneous Contents." Sixteenth ACM Conference on Recommender Systems **(Recsys 2022).**
Top conference in Recommender Systems. *4-page short paper with oral presentation; Acceptance rate: 28.0%.*
Media coverage*: Work was highlighted by [Shaped](#).*

[7] **Wang, Y.**, Zhao, Z., Dai B., Fifty, C., Lin, D., Hong L., Li, W., Chi, E.H.. "Can Small Heads Help? Understanding and Improving Multi-Task Generalization." Proceedings of the ACM Web Conference 2022 (**WWW / theWebConf 2022**).

Full paper with oral presentation; Acceptance rate: 17.7%.

[8] Wang, J., Le, Y., Chang, B., **Wang, Y.**, Chi, E.H., Chen, M.. "Learning to Augment for Casual User Recommendation." Proceedings of the ACM Web Conference 2022 (**WWW / theWebConf 2022**).
Full paper with oral presentation; Acceptance rate: 17.7%.

[9] Oberst, M., D'Amour A., Chen M., **Wang Y.**, Sontag D., Yadlowsky S. Bias-robust Integration of Observational and Experimental Estimators. American Causal Inference Conference (**ACIC 2022**).
Top conference in Causal Inference. ***Journal*** *version on arXiv:* https://arxiv.org/pdf/2205.10467.pdf.

[10] **Wang, Y.**, Wang, X., Beutel, A., Prost, F., Chen, J., Chi, E. H.. "Understanding and Improving Fairness-Accuracy Trade-offs in Multi-Task Learning." Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (**KDD 2021**).
Top conference in Computer Science. *Full paper with oral presentation; Acceptance rate: 15.4%.*
*Short version accepted to BayLearn 2021 (**top** ML conference in the Bay Area).*

[11] Chen, M., **Wang, Y.**, Xu C., Le, Y., Sharma, M., Richardson, L., Wu S., Chi, E.H.. "Values of User Exploration in Recommender Systems." Fifteenth ACM Conference on Recommender Systems (**Recsys 2021**).
Top conference in Recommender Systems. *Full paper with oral presentation; Acceptance rate: 18.4%.*

[12] Chen, Z., **Wang, Y.**, Lin, D., Cheng, D.Z., Hong, L., Chi, E.H., Cui, C.. "Beyond Point Estimate: Inferring Ensemble Prediction Variation from Neuron Activation Strength in Recommender Systems." Proceedings of the 14th ACM International Conference on Web Search and Data Mining (**WSDM 2021**).
Full paper with oral presentation; Acceptance rate: 18.6%.

PATENTS & BLOGS

[13] **Wang, Y.**, Zhang, X., Liu, I., Ning, Y., Peng, C. (2021). "Multi-layer Optimization for a Multi-sided Network Service." **U.S. Patent No. 11,127,066**. Washington, DC: U.S. Patent and Trademark Office.
Patent for work at Uber.

[14] Zhang, X., Zhang, S., **Wang, Y.**, Gogate, M., Ning, Y., Peng, C. Liu, I., Lee, C. (2021). "Optimizing Listing Efficiency and Efficacy for a Delivery Coordination System." **U.S. Patent No. 11,157,579**. Washington, DC: U.S. Patent and Trademark Office.
Patent for work at Uber.

[15] **Wang, Y.**, Ning, Y., Liu, I., Zhang, X. (2018). "Food Discovery with Uber Eats: Recommending for the Marketplace." **Uber Engineering Blog.**
Media coverage*: Selected as* [Top 10 machine learning articles of the month](#) *(0.7%) by an independent publisher.*

WORKING PAPERS

[16] Christakopoulou, K., **Wang, Y.**, Chi, E.H., Chen, M. (2022). "Towards Companion Recommendation Systems."
Submitted to HCAI @ NeurIPS '22 (Human-Centered AI Workshop at NeurIPS).

[17] **Wang, Y.**, Anderson, A. "Impact of Individualism vs. Collectivism on Collective Consumption Behavior on YouTube."
Draft in preparation.

[18] Chang, B., **Wang, Y.**, Chen, M. "Latent Variable Modeling for Consumer Intent Understanding."
Draft in preparation.

[19] Li, P., **Wang, Y.**, Chi, E.H., Chen, M. "Hierarchical Reinforcement Learning for User Exploration Intent in Recommender Systems."
Submitted to theWebConf 2023.

INVITED TALKS & CONFERENCE PRESENTATIONS

"Recommending for a Three-Sided Food Delivery Marketplace: A Multi-Objective Hierarchical Approach." Global Institute for Artificial Intelligence & Business Analytics, Fox School of Business, Temple University, April 2023.

"Recommending for a Three-Sided Food Delivery Marketplace: A Multi-Objective Hierarchical Approach." 2022 Conference on Artificial Intelligence, Machine Learning, and Business Analytics. Harvard University, Boston, MA, December 2022.

"Can Small Heads Help? Understanding and Improving Multi-Task Generalization." Snap Inc TechTalks. November 2022.

"Surrogate for Long-Term User Experience in Recommender Systems." Bay Area Machine Learning Symposium (BayLearn) 2022. South San Francisco, CA, October 2022.

"Recommending for a Three-Sided Food Delivery Marketplace: A Multi-Objective Hierarchical Approach." 2022 INFORMS Annual Meeting. Indianapolis, IN, October 2022.

"Recommending for a Three-Sided Food Delivery Marketplace: A Multi-Objective Hierarchical Approach." Conference on Information Systems and Technology (CIST) 2022. Indianapolis, IN, October 2022 (*Long presentation*).

"Long-Term Planning for Recommender Systems." CONSEQUENCES+REVEAL '22: Causality, Counterfactuals, Sequential Decision-Making & Reinforcement Learning,16th ACM Conference on Recommender Systems (Recsys 2022 Workshop). Seattle, WA, September 2022 (***Invited Speaker and Panelist***).

"Recommending for a Multi-Sided Marketplace with Heterogeneous Contents." 16th ACM Conference on Recommender Systems (Recsys 2022). Seattle, WA, September 2022.

"Moonshot Ally: Assistive Machine Learning for Long-Term User Journeys." Google Brain Summit. Mountain View, CA, September 2022.

"Surrogate for Long-Term User Experience in Recommender Systems." Google Search, Mountain View, CA, September 2022.

"Surrogate for Long-Term User Experience in Recommender Systems." 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington DC, August 2022.

"Understanding and Improving Fairness-Accuracy Trade-offs in Multi-Task Learning." Mays Business School, Texas A&M University. July 2022.

"Recommending for a Three-Sided Food Delivery Marketplace: A Multi-Objective Hierarchical Approach." **ISMS Marketing Science Conference 2022**. Virtual Event, June 2022.

"Understanding and Improving Fairness-Accuracy Trade-offs in Multi-Task Learning." **ISMS Marketing Science Conference 2022**. Virtual Event, June 2022.

"Can Small Heads Help? Understanding and Improving Multi-Task Generalization." ACM The Web Conference 2022. Lyon, France, April 2022.

"Surrogate for Long-Term User Experience in Recommender Systems." Google Research Brain Dump. Virtual Event, February 2022.

"Understanding and Improving Fairness-Accuracy Trade-offs in Multi-Task Learning." Bay Area Machine Learning Symposium (BayLearn) 2021. Virtual Event, October 2021.

"Improving Long-term User Conversion via Surrogate Reward in a REINFORCE Recommender System." Google Research Conference 2021. Virtual Event, October 2021.

"Improving Long-term User Conversion via Surrogate Reward in a REINFORCE Recommender System." Google Research Reinforcement Learning Workshop, Virtual Event, July 2021.

"Understanding and Improving Fairness-Accuracy Trade-offs in Multi-Task Learning." 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Virtual, Singapore, August 2021.

"Understanding and Improving Fairness-Accuracy Trade-offs in Multi-Task Learning." Faire, Data Team. Virtual Meeting, San Francisco, CA, July 2021.

"From User Understanding to User Conversion". Google Research, Brain Team. Mountain View, CA. May 2021.

"Food Discovery with Uber Eats: Holistic Multi-Objective Optimization for the Marketplace." SigOpt. San Francisco, CA, August 2019.

"Food Discovery with Uber Eats: Recommending for the Marketplace." Facebook Research, Core Data Science Team. Menlo Park, CA, June 2019.

"Food Discovery with Uber Eats: Holistic Multi-Objective Optimization for the Marketplace." Airbnb. San Francisco, CA, June 2019.

"Uber Eats Restaurant Ranking and Recommendation." Moving the World with Data Meetup. San Francisco, CA, October 2018.

"Uber Eats Restaurant Ranking and Recommendation." AI Applications @ Uber Eats Meetup. San Francisco, CA, October 2017.

"Robust Approximate Lasso for High-Dimensional Regression." IBM Thomas J. Watson Research Center. Yorktown Heights, NY, February 2016.

"Estimation of High-Dimensional Mean Regression in the Absence of Symmetry and Light-Tail assumptions." Department of Biostatistics, Yale University. New Haven, CT, September 2015.

"Bayesian time series for online query frequency prediction." Internet Services & Research Center, Microsoft Research, Redmond, WA, August 2015.

"Estimation of High-Dimensional Mean Regression in the Absence of Symmetry and Light-Tail assumptions." 2015 Joint Statistical Meetings (JSM), Seattle, WA, August 2015.

"Robust Estimation of High-Dimensional Mean Regression." NSF Workshop for Empr Process and Mod Stat Decision Theory. New Haven, CT, May 2015.

TEACHING EXPERIENCE

Guest Lecture / Tutorials
- Recommender Systems: Algorithms & Applications. *Fuqua School of Business, Duke University.*   *Feb. 2023*
- A Gentle Introduction to Recommender Systems. *Stern School of Business, New York University.*   *June 2022*
- A Gentle Introduction to Recommender Systems. *Heinz College, Carnegie Mellon University.*   *Sept. 2021*
- Experimentation and A/B Testing Best Practices. *Uber Technologies.*   *Oct. 2018*

Mentorship
- Mentor for Undergraduate Consortium at KDD 2022 (KDD-UC)   *June 2022 - present*
- CSRMP mentor for 3 grad & undergrad students from historically marginalized groups   *Nov. 2021 - present*
- Mentor for one **student researcher** and two **interns** at Google Brain   *May 2021 – present*
- Mentor for two **full-time** team members and one **intern** at Uber.   *June 2017 - Sept 2019*

Assistant Instructor at Princeton University
- ORF 504: Financial Econometrics   *Spring 2016*
- ORF 245: Fundamentals of Statistics   *Spring 2014, Spring 2015, Fall 2015*
- ORF 405: Regression and Applied Time Series Analysis   *Fall 2013, Fall 2014*

PROFESSIONAL & ACADEMIC SERVICES

Reviewer: NeurIPS, ICML, CIKM, TheWebConf.   *2019 - Present*

Reviewer: Google PhD Fellowship.	*2021 - Present*
Session chair, Conference on Information Systems and Technology (CIST) 2022.	*June 2022*
Organizer and **session chair**, Workshop on Action, Task and User Journey Modeling.	*Oct 2022*
Session chair, ISMS Marketing Science Conference 2022.	*June 2022*
Session chair, Long-term Dynamics for Responsible Recommendation Systems Workshop'21.	*Nov. 2021*
TPC member, Reinforcement Learning for Real Life (RL4RealLife) Workshop @ ICML 2021.	*July 2021*

SELECTED AWARDS

Cummins Merit Fellowship, Princeton University	*Jan. 2013*
Guo Moruo Scholarship, USTC	*May 2012*
Outstanding Research Award, USTC	*Feb. 2012*
CSST Award, UCLA	*Sept. 2011*
National Scholarship, Ministry of Education of China	*Nov. 2010, Nov. 2009*

SKILLS

Programming skills: Python (proficient), R (proficient), TensorFlow (proficient), Hive, SQL, Spark, C/C++, Matlab, q/kdb+

Languages: Mandarin (native), English (fluent)

OTHERS

Brain STAR Women Steering Committee at Google Brain	*2022 – present*
Member of Women Lean-in Circle at Google Brain	*2020 - present*
Organizer of Reinforcement Learning Research & Engagement Team Meeting at Google Brain	*2022 - present*
Organizer of "Assistive ML for Long-Term User Journeys" Moonshot biweekly meeting	*2021 – present*
Member of WiSDOM (Women in Statistics, Data, Optimization and Machine Learning) at Uber	*2018 - 2019*
Organizer of Discovery team Lunch & Learn, Eater Data Science reading group at Uber	*2017 - 2019*
Organizer of Wilks Statistics Seminar at Department of ORFE at Princeton University	*2015 - 2016*
Vice President of Student Union of School for Gifted Young at USTC	*2011 - 2012*