# Myungha Jang

*Applied Scientist / Machine Learning Engineer*

☎ +1-413-992-3413
✉ myunghajang@gmail.com
myungha.me

## Profile

I am a self-motivated machine learning engineer with research background in information retrieval, graph mining, and NLP. I am passionate about solving user-facing problems and skilled in setting a scientific direction to make improvements in ranking and recommendation systems

## Education

2013 - 2019 **Ph.D. in Computer Science**, *University of Massachusetts Amherst*, MA, U.S., Center for Intelligent Information Retrieval.
Ph.D. Thesis: Probabilistic Models for Identifying and Explaining Controversy

2011-2012 **M.S. in Computer Science**, *POSTECH*, Pohang, South Korea.

2006-2011 **B.S. in Computer Science**, *Ewha Womans University*, Seoul, South Korea.
Study abroad at Wesleyan College, U.S. in 2009
Graduated with Magna Cum Laude

## Experience

Feb. 2022 - Jan 2023 **Senior Machine Learning Engineer**, *Meta*, Menlo Park, CA.

- **User Taste Graph in Applied Research (July 2022 - Jan 2023)**: User Taste Graph is an applied research team that develops graph-related techniques, such as graph clustering, KNN in graphs, GNN. To assist other product ranking teams in addressing their challenges. We collaborated with the Instagram Reels recommendation team to enhance their retrieval phase by identifying more relevant video candidates. I led several workstreams aimed at extracting better topical interests of users. For example, one workstream involved testing a modified personalized page rank algorithm, another one integrating long-term and short-term interests.
- **Facebook Groups Content Quality (July 2021 - June 2022)**: As a senior ML engineer, I led a cross-functional team of 10 people, including 4 production engineers, 1 PM, 3 ML engineers (including myself), 2 EM, and 1 data scientist, with the goal of giving users negative feedback control. We launched a UI for users to provide negative feedback (e.g., x-out button, âĂIJhide postâĂİ button), and built ML models to predict usersâĂŹ actions, demoting them from group post ranking. Additionally, I wrote a roadmap to specify goals and align milestones and project deadlines with different parties (ML ranking and the product team). I launched two personalized ML models in production for the group post ranking system. These models predict the probability of a given content being hidden by users and demote such content in the final ranking.

Feb. 2019 - Feb. 2022 **Machine Learning Engineer** , *Meta*, Menlo Park, CA.

- **'Interest Pages to Follow' Recommendation (Jan 2021 - July 2021)** I collaborated with two other ML engineers to build Facebook's new "interest" page recommendation system from a very early phase. During the retrieval phase, I implemented and tested several KNN-based candidate generators to expand our inventory 150K to 3M pages. For the ranking phase, I improved our GDBT ranking model by developing a deep learning model that resulted in increased click rates in online experiments.
- **Facebook Hashtag Search Ranking (Feb 2020 - Jan 2021)**: I implemented several ranking features in C++ to increase query relevance and led the workstream focused on ensuring the integrity of election-related hashtags during the 2020 Presidential Election period. We automatically identified growing election-related hashtags and tested a stricter filter for ranking to address integrity issues.
- **Personalized Integrity (Feb 2019 - Feb 2020)**: I built personalized ML models that predicts users' preference against subjective borderline-bad content on Facebook Newsfeed and created metrics that study polarization of Newsfeed with collaboration with social scientists.

Summer 2017 **Machine Learning Software Engineer Intern**, *Facebook*, Menlo Park, CA.

- Machine Learning Intern in Whole Page Ranking Team
- Conducted various search A/B testings and analysis and worked on developing search results ranker

Summer 2015 **Research Intern**, *IBM T.J. Watson Research Center*, Yorktown heights, NY.

- Host: Dr. Kenneth W. Church
- Worked on improving Watson's Automatic QA Retrieval System for customer questions regarding IBM products by analyzing and enhancing the quality of the training dataset using various NLP techniques

2013- Present **Research Assistant**, *Center for Inteligent Information Retrieval*, UMass Amherst.

- My research topics broadly lied on understanding controversial search query and results to help users with critical literacy. My research specifically focused on building probabilistic models to find controversial topics in the documents and create automatic summary to explain why the topic is controversial by finding representative opinions of conflicting stances from social media posts
- Advised by Prof. James Allan

2011-2012 **Research Assistant**, *Information and Database Systems Lab*, POSTECH.

- Conducted research on inferring latent comparable entity relationship from query log graph to provide users a complete suggestion of other comparable entities in search query
- Advised by Prof. Seung-won Hwang

## Selected Publications

### Conferences

- Youngwoo Kim, **Myungha Jang** and James Allan, Explaining Text Matching on Neural Natural Language Inference, ACM Transactions on Information Systems (ToIS) 2021
- **Myungha Jang** and James Allan, Explaining Controversy on the Social Media via Stance Summarization, The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018)
- **Myungha Jang**, Shiri Dori-Hacohen and James Allan, Modeling Controversy Within Populations, The 3rd ACM International Conference on the Theory of Information Retrieval (ICTIR 2017)
- **Myungha Jang**, Jinho Choi, James Allan, Improving Document Clustering by Removing Unnatural Language, The 3rd Workshop on Noisy User-generated Text at EMNLP (W-NUT 2017)
- **Myungha Jang**, John Foley, Shiri Dori-Hacohen, and James Allan, Probabilistic Approaches to

Controversy Detection, The 25th ACM International Conference on Information and Knowledge Management (CIKM 2016)

- **Myungha Jang**, James Allan, Improving Automated Controversy Detection on the Web, The Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2016) pp. 865-868
- Jongwuk Lee, **Myungha Jang**, Dongwon Lee, Won-Seok Hwang, Jiwon Hong, Sang-Wook Kim, Alleviating the Sparsity in Collaborative Filtering using Crowdsourcing, *The Int'l Workshop on Crowdsourcing and Human Computation for Recommender Systems (CrowdRec)*, 2013.
- **Myungha Jang**, Jin-woo Park, Seung-won Hwang, Predictive Mining of Comparable Entities from the Web, *The 26th Association for the Advancement of Artificial Intelligence (AAAI)*, pp. 66-72, 2012.

## Selected Awards

2016 KSEA-KOCSEA Graduate Scholarship

2016 Best Poster Award at UKC 2016

2014 Grace Hopper Celebration Scholarship

2009 2010 Undergraduate Research Program Fellowship ($10,000), Korea Foundation for the Advancement of Science and Creativity

## Activities

**CS Women Group Co-chair at UMass Amherst** 2014–2015

I served to organize programs and services for female students in the CS department at UMass.

**Ewhaian (*www.ewhaian.com*) Administrator** 2006–2008

Ewhaian is the biggest online community for students of Ewha Womans University with approximately 60,000 users. I managed the web server and the web design, helped to organize four large fund-raising parties of 5,000 people, and designed on/offline posters for numerous events.

## Technical Skills

Programming: Java, Python, Pytorch, Tensorflow, C/C++, AWK/Bash, SQL
OS: Windows, UNIX, MAC OS
Tools: Adobe Illustrator and Photoshop