

# Gang Cheng

210 Talbert Hall, Buffalo, NY 14260

☎ (716)-907-6725

✉ withchenggang@gmail.com

in linkedin.com/in/withchenggang

## SUMMARY

---

**Machine learning engineer** with 4+ years of work experience in a range of Search & Recommendations projects. Developed and deployed large-scale machine learning models to serve **200M+** users at Weibo.

## EDUCATION

---

**University at Buffalo, The State University of New York**

*M.S. in Computer Science and Engineer*

Buffalo, NY

Aug. 2021 - May 2023

**Shanghai University of International Business and Economics**

*B.S. in Applied Statistics*

Shanghai, China

Sept. 2013 - Jul. 2017

## EXPERIENCE

---

**Weibo Corporation**

*Machine Learning Engineer*

Beijing, China

Aug. 2020 - Jul. 2021

- Designed and developed an **Embedding-based Retrieval** framework for candidate generation at Weibo machine learning platform for personalized recommendations for **200M+** users every day at Weibo
- Built **Graph Convolutional Network** algorithm on a graph with **millions of nodes and billions of edges** to generate embedding of item for item-item recommendation task

**Qutoutiao Inc.**

*Machine Learning Engineer*

Shanghai, China

Nov. 2018 – May 2020

- Improved CTR by **4.8%** using deep learning **ranking** models with **large-scale sparse features** on the task of personalized recommendations of news feed
- Developed the **two-tower DNN model** to retrieve top relevant candidates accurately and efficiently leading to **12%** increase of CTR of news feed
- Built large-scale personalized **recommendation engines** using **Golang** for data pipelines for feature generation, training examples generation at Qutoutiao

**Trip.com Group**

*Machine Learning Engineer*

Shanghai, China

Nov. 2017 – Nov. 2018

- Trained the **Matrix Factorization** model using **Scala** with **Spark MLlib** package for personalized retrieval
- Optimized the **LambdaMART** algorithm using **Xgboost** and improved the Top-3 recommendation accuracy by **10%** in Trip.com's hotel search task

## PROJECTS

---

**Embedding-based Retrieval Framework**

Aug. 2020 - Apr. 2021

- **Distributed training:** Developed DNN models with various neural network architectures using **Tensorflow**
- **Model serving:** Deployed models using **TensorFlow Serving** to calculate user embedding in real time
- **Offline inference:** Generated training examples to calculate item embedding every half hour
- **Similarity search:** Developed Weiss based on **Faiss** to search top candidates with **ANN algorithms**

## SKILLS

---

- **Languages:** Python, Golang, Java, Scala, R, SQL, Shell
- **Frameworks & Tools:** Tensorflow, Scikit-learn, PySpark/Spark, AWS, Kubernetes, Mysql, Redis, Kafka, Git
- **Machine Learning:** Profound knowledge in Applied Machine Learning, Large-Scale Recommendation System, Graph Algorithms and Natural Language Processing