

# Lei Zuo

(646)-203-6488 | lzuo98@gmail.com | leonardzuo.com | 435 West 119th Street, New York, NY 10027

## EDUCATION

---

### Columbia University

M.S. in Computer Science, GPA: 3.78

New York  
Exp Dec 2022

- Selected Coursework: Large-Scale Distributed Systems, Engineering Software-as-a-Service, Conversational AI

### The University of Hong Kong

B.Eng. (Hons.) in Computer Science

Hong Kong  
Jun 2020

- Selected Coursework: Software Engineering, Operating Systems, Database, Functional Programming

## PROFESSIONAL EXPERIENCE

---

### TikTok

Software Engineering Intern

Mountain View  
May 2022 – Aug 2022

- Worked on high-performance multi-modality data loader in Data I/O Team, Applied Machine Learning Department
- Built caching functionality for HDFS data with PyArrow, reducing runtime by 30% for 10+ end2end training tasks
- Implemented triplet sampling on KV datasets with Python, achieved parallel computing by sharding data among different GPUs, solved scaling issues by generating and querying metadata
- Developed and debugged packing scripts to convert 10+ KV datasets into parquet format
- Wrote unit tests for all developed features, provided detailed documentation to clients

### Fovea

Machine Learning Engineering Intern

New York  
Jul 2021 – Aug 2021

- Designed and implemented a self-supervised joint visual-linguistic pre-trained model with TensorFlow
- Synthesized video description, achieved 12%/7% improvement on BLEU-3/4 score over baseline

### Axisplay

Software Engineering Intern

Hong Kong  
May 2020 – Aug 2020

- Implemented 10+ features (production) using React.js and Kotlin for JsBridge-based Android app
- Worked closely with QA in the agile development cycle, troubleshooting 50+ issues for on-time releases

### Bank of China

Information Technology Intern

Hong Kong  
Jun 2018 – Aug 2018

- Automated sentiment analysis on customer reviews scraped daily from multiple websites with scikit-learn
- Performed EDA and cleaning on FX records with Python to help the team build models that resulted in 4X less MSE

## PROJECTS

---

### Novel Summarization with Graph Neural Network

May 2021 – Present

- Designed data augmentation scheme and few-shot models for domain adaptation of summarization system
- Benchmarked the task by building graph node alignment classifier with both word2vec and BERT embeddings

### Distributed Systems

Feb 2022 – May 2022

- Built a key/value storage system with Golang that shards keys over a set of replica groups for parallel operation
- Developed a Paxos-based key/value storage system and a model checker for the correctness of Paxos state machine

### Course Management SaaS

Aug 2021 – Jan 2022

- Utilized Ruby on Rails to build a course management website with extra functionalities
- Developed unit tests for the SaaS product with Cucumber and RSpec, achieving 90%+ coverage

### Dialog System for All

Sep 2021 – Dec 2021

- Built a task-oriented multilingual chatbot with PyTorch for 8 languages

## PUBLICATIONS

---

Improving Conversational Recommendation System's Quality with Context-Aware Meta Information

NAACL 2022

- Built conversational recommendation system with PyTorch, increasing recall@1 by 40% and BLEU-4 by 30%

Comparing Integer Linear Programming to SAT-Solving for Hard Problems...

AICoB 2020

- Developed Python pipelines that solve computational problems with SAT solvers, reducing runtime by up to 700X

## TECHNICAL SKILLS

---

**Languages:** Python, C/C++, Java, JavaScript, HTML/CSS, Haskell, Unix Shell, R, Ruby, Kotlin, Golang

**Frameworks:** PyTorch, scikit-learn, TensorFlow, PyArrow, pandas, OpenCV, React.js, Node.js, Rails, RSpec, Cucumber

**Toolkits:** Firebase, Google Cloud Platform, MongoDB, SQLite, webpack, Docker, Anaconda, Git, Hadoop