Taeksang Peter Kim

tspeterkim3@gmail.com | 582-263-1348 | tspeterkim.github.io

EDUCATION

UIUC
MENG in Electrical and Computer Engineering
Dec 2025 | Champaign, IL

UCLA
BS + MS in Computer Science
Jun 2017 + Dec 2018 | Los Angeles, CA
Cum. GPA: 3.7 / 4.0

INTL. SCHOOL OF BRUSSELS
Valedictorian
June 2013 | Brussels, Belgium

LINKS

Github:// **tspeterkim**
LinkedIn:// **tspeterkim3**
X:// **@tspeterkim**

PROJECTS

MINIMAL FLASH ATTENTION
Flash Attention in ~100 lines of CUDA
400+ Stars on Github
Made the **front page** of Hacker News

ONE BILLION ROW CHALLENGE IN CUDA
First CUDA solution to the challenge
Runs in 17s on a V100
Made the **front page** of Hacker News

SKILLS

PROGRAMMING
Python • Go • Java • SQL • Shell

TOOLS
AWS • PyTorch • TensorFlow • Pandas
Ray • Spark • Jenkins • Airflow

EXPERIENCE

COUPANG | Senior ML Engineer
Aug 2022 - Aug 2024 | Seoul, South Korea
- Built the recommendation system for the largest South Korean streaming service by MAU.
- Used GPT to do next watch prediction and productionized it as the default re-ranking algorithm, raising overall playback time by 20%.
- Applied Thompson Sampling to deliver artwork personalization, raising feed engagement by 14%.

BUZZVIL | ML Engineer
Mar 2019 – Jul 2022 | Seoul, South Korea
- Improved e-commerce ads for Buzzvil, South Korea's 3rd largest mobile advertising platform (after Google and Kakao).
- Increased ad spending return by 30% through a user-action sequence based BERT model and an Item CF model.
- Designed a company-wide introductory ML course, and made it the most popular study in the company (15% of the entire company enrolled).

LINKEDIN | ML Engineering Intern
Jun 2018 – Sep 2018 | Mountain View, CA
- Trained and fine-tuned job response prediction models for the search page that scaled up to 500M+ members.
- Performed feature distribution analyses with Apache Pig scripts.

PAYPAL | Software Engineering Intern
Jun 2016 – Sep 2016 | San Jose, CA
- Deployed Apache Eagle on top of existing Hadoop clusters as the first real-time log monitoring tool deployed in production.
- Contributed to the Apache Eagle project by future-proofing a HBase log parser.

RESEARCH

UCLA NLP LAB | Research Assistant
Sep 2017 – Dec 2018 | Los Angeles, CA
Worked with **Wasi Ahmad** and **Prof Kai-Wei Chang** to build a sentence-based document retriever using Apache Lucene for open-domain SQuAD, and implement the first TensorFlow version of the Stanford Attentive Reader.

PATENTS

[1] Taeksang Kim. Personalized video DJ with text-to-video, Korean Pat. No. 10-0123761, 2023.

[2] Taeksang Kim. Video keyword tagging with large language models, Korean Pat. No. 10-0098584, 2023.

[3] Jina Hwang and Taeksang Kim. Sports awareness, mobile feed video player, Korean Pat. No. 10-0071581, 2023.

[4] Taeksang Kim. Next watch prediction using GPT, Korean Pat. No. 10-0036443, 2023.

[5] Taeksang Kim. First hero optimization through reinforcement learning, Korean Pat. No. 10-0184572, 2022.