

AI startup Upstage has achieved the best performance in Korea in data-centric research.

Upstage (CEO Kim Sung-hoon, www.upstage.ai) announced that it has achieved the highest research performance among domestic companies by presenting seven papers at the Data-centric Machine Learning Research (DMLR) Workshop, the most prestigious workshop in the field of Data-Centric AI.

ICML 2023-DMLR is the highest level of academic event dealing with the theory and practice of machine learning, and it is the most influential artificial intelligence conference with a paper adoption rate of 20% and an impact factor of 6.99, with about 70,000 people participating every year.

The DMLR, where Upstage's data team presented its research results, is a workshop on data-centric machine learning research (Data-centric Machine Learning Research) among ICML 2023, and it is a follow-up workshop to the NeurIPS 2021- Data Centric AI workshop, which was created by Stanford University Professor Andrew Ng, a world-renowned scholar and founder of Data-Centric AI.

This year's DMLR will be the third largest research exchange in the field of Data-Centric AI, following the NeurIPS 2021- Data Centric AI workshop and the ICML2022 Dataperf workshop last year. It will be held in Hawaii on July 29, 2023.

Upstage will present a total of seven research results at the workshop, including "DMOps: Data Management Operations and Recipes," which proposes a process for effectively building artificial intelligence learning data in companies, and "Inter-Annotator Agreement in the Wild: Uncovering Its Emerging Roles and Considerations in Real-World Scenarios," which enables effective cost management and efficient construction when building data.

Among this year's research results, four papers were jointly researched with the research team of

Professor Lim Hee-seok of Korea University, a world-renowned scholar in the field of natural language processing, and it is evaluated that it opened a new horizon for joint research between startups and academia.

Through joint research, Upstage and Korea University have achieved various results, including data-centric machine translation, data-centric common sense reasoning research, and a paper analyzing the negative effects of synthetic data, as well as proposing a new methodology for the post-processing field of speech recognition, "Explainable Error Benchmark."

In particular, Upstage researcher Park Chan-joon, who participated in all seven of Upstage's papers, has presented papers in this field for three consecutive years since 2021, contributing to Upstage's development as a global "Data-Centric AI" specialist company.

With this research result, Upstage has reaffirmed its global top-level technology in the field of Data-Centric AI, from basic research to applied research. Upstage has established itself as an AI top technology startup by presenting papers at all three major AI conferences, including ICLR, CVPR, and ECCV, last year.

In addition, it attracted attention by ranking in the top five in Korea, the only startup among Samsung, Naver, Kakao, and LG, in the '2021 Global AI Top Tier Conference Paper' survey released by the UK's 'Imperial College London' last year.

Upstage CEO Kim Sung-hoon said, "I am very pleased to be able to present global top-level research results at this year's ICML 2023-DMLR." "Upstage will do its best to make everyone use the best performance AI more conveniently through not only AI services but also generous R&D investment."