

# JAEWON CHEON

## RESEARCH INTERESTS

Large Language Models, Efficient Sequence Modeling

## EDUCATION

<b>Korea University</b> <i>M.S., Industrial and Management Engineering (Advisor: Prof. Pilsung Kang)</i>	Mar 2024 - Feb 2026
· Thesis: Leveraging Intrinsic Sparsity of GLU-based Large Language Models for Inference Efficiency	GPA: 4.5/4.5
<b>Korea University</b> <i>Bachelor of Media and Communication</i>	Mar 2016 - Feb 2023
	GPA: 4.38/4.5

## WORK EXPERIENCE

<b>Naver</b>   <i>Foundation Model Team</i>	Jan 2026 - Apr 2026
· Research Internship	
· Topic: LLM Reasoning, Efficient Learning	
<b>Upstage</b>   <i>NAVER Connect BoostCamp AI Tech</i>	Sep 2025 - Nov 2026
· Teaching Assistant	
· Topic: LLM, Generative AI	
<b>Tesser Inc.</b>   <i>Natural Language Processing Team</i>	Oct 2023 - Jan 2024
· Research Internship	
· Topic: Keyword extraction, Medical LLM	
· Contributed to <a href="#">Ontol</a> @ Tesser	

## PUBLICATIONS

- [1] [COUNTDOWN: Contextually Sparse Activation Filtering Out Unnecessary Weights in Down Projection](#)  
Jaewon Cheon and Pilsung Kang.  
*Empirical Methods in Natural Language Processing (EMNLP) Main Track, 2025*

## AWARDS & GRANTS

<b>Basic Science Research Program</b>   <i>Granted by National Research Foundation of Korea</i>	Sep 2024 - Aug 2025
· Inference Acceleration of Large Language Models via Sparsity Prediction	

## CONFERENCE PRESENTATIONS

- [Enhancing Manufacturing Reliability Through NLP-Based Log Anomaly Detection](#)  
· Presented at the *2025 INFORMS Conference on Quality, Statistics and Reliability (ICQSR), Singapore*

## PROJECTS

---

<b>Train Fast &amp; Infer Fast</b>   <i>Ministry of Science and ICT</i>	Sep 2025 - Feb 2026
· Implemented a self-distillation approach for transforming full-attention LLMs into hybrid RNN–attention models	
<b>Generative AI-based Workflow Automation</b>   <i>Samsung Fire &amp; Marine Insurance</i>	Apr 2025 - Dec 2025
· Built hierarchical data conversion pipeline for insurance policy documents for RAG-based workflow automation.	
<b>Researcher Matching System</b>   <i>Seoul National University</i>	Dec 2024 - Jun 2025
· Developed a RAG-based recommendation system that matches industry partners with suitable researchers	
<b>Development of a News-Domain Specialized RAG Pipeline</b>   <i>BECUAI</i>	Apr 2024 - Jun 2024
· Enhancing search quality and answer reliability through query refinement and context-aware response structuring	