Taehun Cha

cth127@korea.ac.kr · https://github.com/cth127

Taehun Cha is a Ph.D. candidate at the Department of Mathematics, Korea University. His main research area is Natural Language Processing and sequential decision-making. As an intersection, he is currently interested in Generative Multi-Agent Systems. He is also interested in mathematically analyzing the current success of PLMs and LLMs.

Education

Ph.D. Candidate in Mathematics – Korea University — Mar. 2022 —

M.F.E. in Financial Engineering – Korea University — Mar. 2020 — Feb. 2022

- Fully funded student
- Academic Excellence Scholarship for 2021 Spring Semester
- Thesis: Understanding the Yield Curve Shift with FOMC Statements: NLP Perspective

B.A. in Sociology and Cultural Critics – Yonsei University — Mar. 2012 — Aug. 2019

- Minor in Applied Statistics

Work Experience

KT Corp., Ph.D. Student Researcher, Jul. 2023. - Aug. 2023.

- Researched the hallucination problem in large language models (LLM). Built an automatized pipeline to construct a hallucination dataset using ChatGPT and a reward model to train LLM with RL.
- Selected as an outstanding intern.

Publication

Taehun Cha and Donghun Lee. 2024. “Pre-trained Language Models Return Distinguishable Probability Distributions to Unfaithfully Hallucinated Texts.” Findings of the Association for Computational Linguistics: EMNLP 2024 **(EMNLP 2024 Findings)**.

Taehun Cha and Donghun Lee. 2024. “Evaluating Extrapolation Ability of Large Language Model in Chemical Domain.” Language + Molecules Workshop at ACL 2024. **(Lang+Moles@ACL 2024)**

Taehun Cha and Donghun Lee. 2024. “SentenceLDA: Discriminative and Robust Document Representation with Sentence Level Topic Model.” In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics **(EACL 2024, Oral)**.

Taehun Cha and Donghun Lee. 2023. "Predicting U.S. Treasury Yield Curve Shifts with FOMC Statements Using BERT." In Proceedings of the 50th Korea Computer Congress **(KCC 2023)**.

Taehun Cha*, Jaeheun Jung*, and Donghun Lee. 2022. “Noun-MWP: Math Word Problems Meet Noun Answers.” In Proceedings of the 29th International Conference on Computational Linguistics **(COLING 2022)**.

Awards

AI Grand Challenge - 7nd Place. 2023. Ministry of Science and ICT.

- Designed and led the development of an open-domain, multi-hop, multi-modal, document-based report-generating system (team leader).

AI Grand Challenge Open Track - 2nd Place. 2023. Ministry of Science and ICT.

- Developed an open-domain, multi-hop, multi-modal document QA system and achieved 2nd place out of 12 teams.

Korean AI Competition - 4th Place. 2022. Ministry of Science and ICT.

- Developed an Automatic Speech Recognition model for the Korean language and achieved 4th place out of 103 teams (team leader).

AI Grand Challenge. 2021. Ministry of Science and ICT.

- Developed an NLP model to solve elementary math word problems based on KLUE-RoBERTa, and we were selected for follow-up research (team leader).

HAAFOR NLP Challenge 2020 - 3rd place. 2020. HAAFOR.

- Achieved 3rd place(70.27% accuracy) on text order prediction task with ALBERT model.

K-Cyber Security Challenge 2019 - final round. 2019. KISA.

- Developed a random forest algorithm to predict and detect a cyber attack on automobiles.

Academic Service

- Reviewer for ACL Rolling Review
- Reviewer for Neural Computing and Applications

Grant

Junior Fellow-Research Grant. 2023. Korea University.

- Financial support for research on sentence-level topic modeling.

Invited Talks

Extracting Financial Knowledge from FOMC Communications with Reinforcement Learning and Natural Language Processing, The Artificial Intelligence Symposium, Natural Science Research Institute, Gangneung-Wonju National University, June 9, 2023.