Ercong Nie

✉ ecnie@foxmail.com   🎂 1997.04.21
📞 +49 15208326987   🏠 Oettingenstraße 67 D-80538 Munich, Germany
in Ercong Nie   👤 Personal Page: `https://www.cis.lmu.de/~nie`

🎓 EDUCATION

Ludwig-Maximilians-University of Munich (LMU), Munich, Germany   2022.09 – present

PhD student in Natural Language Processing (NLP) at Center for Information and Language Processing (CIS), supervised by Dr. Helmut Schmid and Prof. Hinrich Schütze
Research fields: *multilingualality of large language models, low-resource NLP, human-centered NLP, etc.*

Ludwig-Maximilians-University of Munich (LMU), Munich, Germany   2020.04 – 2022.08

M.Sc. in Computational Linguistics (Major) and Computer Science (Minor)   Grade: 1.3/1.0
Selected courses: *Python Applications for Computational Linguistics, Neural representation learning, Machine translation, Conversational AI, Deep learning for NLP, Machine learning, Data mining, Big data etc.*
Master Thesis: *Zero-Shot Learning Transfer on Low-Resource Languages by Cross-Lingual Retrieval* (Examiner: Prof. Hinrich Schütze)

Shanghai Jiao Tong University (SJTU), Shanghai, China   2015.09 – 2019.06

B.A. in German Studies (Major) and Finance (Minor)   Grade: 88/100
Selected courses: *General Linguistics, German Lexicology, Neurolinguistics, Mathematics, Statistics etc.*
Bachelor Thesis: *A corpus-based study of metaphors in the football language* (Supervisor: Prof. H. Zhang)

Ruprecht-Karls-University Heidelberg (RKU), Heidelberg, Germany   2018.03 – 2018.08

 Exchange Student in Comparative German Studies
Study courses: *Language typology, Psycolinguistics, Second Language Acquisition, etc.*

👥 EXPERIENCE

Doctoral Researcher in Natural Language Processing, Munich, Germany   2022.09 – present

Doctoral Researcher   Working at Schuetze NLP Lab of LMU Munich and MCML (Munich Center for Machine Learning)

Computational Linguistic Internship at Infineon, Munich, Germany   2021.08 – 2022.03

Working Student   Member of Semantic Web Group, Supply Chain Innovation Team

- Performed ontology modeling and linguistic analysis of ontologies and knowledge graphs.
- Developed a document retrieval system for internal use by employing the language model sentence BERT and the ontology digital reference.

Computational Linguistic Research Assistant at CIS, Munich, Germany   2021.04 – 2021.07

Student Research Assistant   DFG Project ReMLAV: Relational Machine Learning for Argument Validation

- Worked on the construction of a corpus for claim detection.
- Annotated argument sentences by claim target, claim stance and claim type for argument classification tasks.

📂 PUBLICATIONS

⋆ Equal contribution, † Corresponding author.

1. **ToPro: Token-Level Prompt Decomposition for Cross-Lingual Sequence Labeling Tasks.** *EACL 2024*. Bolei Ma*, Ercong Nie*, Shuzhou Yuan, Färber Michael, Frauke Kreuter, Helmut Schmid, and Hinrich Schütze.

2. **Language Models as the Brain: Probing Linguistic Structures within Neural Language Models using Minimal Pair.** *LREC-COLING 2024*.
Linyang He, Peilin Chen, Ercong Nie, Yuanning Li, and Jonathan R. Brennane.

3. **Decomposed Prompting: Unveiling Multilingual Linguistic Structure Knowledge in English-Centric Large Language Models.** *Preprint 2024*.
Ercong Nie, Shuzhou Yuan, Bolei Ma, Helmut Schmid, Michael Färber, and Frauke Kreuter, Hinrich Schütze.

4. **GNNavi: Navigating the Information Flow in Large Language Models by Graph Neural Network.** *Preprint 2024*.
Shuzhou Yuan, Ercong Nie, Bolei Ma, Michael Färber, Helmut Schmid, and Hinrich Schütze.

5. **Why Lift so Heavy? Slimming Large Language Models by Cutting Off the Layers.** *Preprint 2024*.
Shuzhou Yuan*, Ercong Nie*, Bolei Ma, and Michael Färber.

6. **Unleashing the Multilingual Encoder Potential: Boosting Zero-Shot Performance via Probability Calibration.** *Findings of EMNLP 2023*.
Ercong Nie, Helmut Schmid, and Hinrich Schütze.

7. **Cross-Lingual Retrieval Augmented Prompt for Low-Resource Languages.** *Findings of ACL 2023*.
Ercong Nie*, Sheng Liang*, Helmut Schmid, and Hinrich Schütze.

8. **From Classification to Generation: Insights into Crosslingual Retrieval Augmented ICL.** *NeurIPS 2023 Workshop on Instruction Tuning and Instruction Following*.
Xiaoqian Li, Ercong Nie and Sheng Liang.

9. **Baby's CoThought: Leveraging Large Language Models for Enhanced Reasoning in Compact Models.** *EMNLP 2023 Workshop CoNLL-CMCL Shared Task BabyLM Challenge*.
Zheyu Zhang*, Han Yang*, Bolei Ma*, David Rügamer and Ercong Nie[†].

10. **Crosslingual Retrieval Augmented In-context Learning for Bangla.** *EMNLP 2023 Workshop on Bangla Language Processing (BLP 2023)*.
Xiaoqian Li, Ercong Nie and Sheng Liang.

11. **Is Prompt-Based Finetuning Always Better than Vanilla Finetuning? Insights from Cross-Lingual Language Understanding.** *KONVENS 2023*.
Bolei Ma*, Ercong Nie*, Helmut Schmid, and Hinrich Schütze.

12. **Cross-Lingual Constituency Parsing for Middle High German: A Delexicalized Approach.** *RANLP 2023 Workshop on Ancient Language Processing (ALP 2023)*.
Ercong Nie, Helmut Schmid, and Hinrich Schütze.

13. **What cleaves? Is proteasomal cleavage prediction reaching a ceiling?** *NeurIPS 2022 Workshop on Learning Meaningful Representations of Life (LMRL 2022)*.
Ingo Ziegler, Bolei Ma, Ercong Nie, Bernd Bischl, David Rügamer, Benjamin Schubert, and Emilio Dorigatti.

⇄ TEACHING

Courses

- **Human-centric Natural Language Processing Seminar** (Master). *Teaching Assistant*.
  CIS, LMU Munich   SS 2024 & SS2023
- **Statistical Methods in the Language Processing** (Bachelor). *Teaching Assistant*.
  CIS, LMU Munich   WS 2023-24
- **Foundations of Advanced Natural Language Processing** (Master). *Teaching Assistant*.
  CIS, LMU Munich   WS 2023-24 & WS 2022-23
- **Applied Deep Learning** (Master). *Supervisor*.
  Department of Statistics, LMU Munich   WS 2023-24 & SS 2023
- **Finite State Technologies** (Master) *Teaching Assistant*.
  CIS, LMU Munich   WS 2021-22

Thesis Supervision

- **Enhancing Retrieval Augmented Generation (RAG) for Domain-Specific Content** *(with Bosch)*
  Master Thesis, CIS, LMU Munich   SS 2024
- **Enhancing Reasoning and Safety: Integrating Classical Rule-Based AI with Large Language Models**

Master Thesis, CIS, LMU Munich   WS 2023-24
- **Prompt-based finetuning of multilingual models for zero-shot cross-lingual transfer**
Master Thesis, CIS, LMU Munich   SS 2023
- **Extend the cross-lingual retrieval-augmented prompting method to new tasks and settings**
Master Thesis, CIS, LMU Munich   SS 2023

Student Project Supervision

- **SemEval'24 Task 1: Measuring the Semantic Textual Relatedness.**
Course project, Department of Statistics, LMU Munich   WS 2023-24
- **SemEval'24 Task 8: Machine-Generated Text Detection**
Course project, Department of Statistics, LMU Munich   WS 2023-24
- **BabyLM Challenge: Sample-efficient pretraining on a developmentally plausible corpus**
Course project, Department of Statistics, LMU Munich   SS 2023

⚙ Skills

- Programming Languages: Python (Good), C++ (Basic), familiar with common NLP and machine learning packages, such as Numpy, Scipy, Scikit-Learn, PyTorch, Keras, NLTK, etc.
- Natural Languages: Chinese Mandarin (Native), English (Fluent), German (Fluent)
- Tools: LaTeX, HTML, Linux, Jupyter Notebook, Git, etc.

♡ Honors and Certificates

Outstanding Graduate of Shanghai Jiao Tong University	2019
Academic Excellence Scholarship of Shanghai Jiao Tong University	2016, 2017, 2018
Merit Student of Shanghai Jiao Tong University	2016, 2018

ℹ Others

Invited talk at 31st Student Conference on Computational Linguistics (TaCoS): *Fine-Tuned Sentence Transformer Model for Question Answering Task*   2022.05