Yuwei Yin (Joey) 阴昱为

Homepage ‖ Scholar ‖ GitHub ‖ LinkedIn ‖ Email: seckexyin@gmail.com

EDUCATION

- **The University of British Columbia (UBC)** — Vancouver, BC, Canada
  M.Sc. in Computer Science (Ph.D. Track) at Department of Computer Science — *Jan. 2024 – Present*

- **Beihang University (BUAA)** — Haidian, Beijing, China
  ~~*Ph.D.*~~ *(Dropped-out)* in Computer Science at State Key Lab of Software Dev. Env. — *Sep. 2019 – Sep. 2021*
  ~~*M.Eng.*~~ *(Transferred)* in Computer Technology at School of Computer Sci. & Eng. — *Sep. 2018 – Jun. 2019*
  GPA: 92/100; Rank 3%

- **Beijing Information Science and Technology University (BISTU)** — Chaoyang, Beijing, China
  B.Eng. in Computer Science and Technology at School of Computer Science — *Sep. 2014 – Jun. 2018*
  GPA: 93/100; Rank 1st of 202

EMPLOYMENT

- **The University of British Columbia (UBC)** — Vancouver, BC, Canada
  <u>*Teaching Assistant*</u> *(Full-time, Onsite, Paid)* — *Jan. 2024 – Apr. 2024*
    - **Affiliation**: Department of Computer Science, Faculty of Science
    - **Duty**: Assisting professors in teaching CPSC 440/550 Advanced Machine Learning.

- **The University of Hong Kong (HKU)** — Pokfulam, Hong Kong
  <u>*Research Assistant*</u> *(Full-time, Onsite, Paid)* — *Oct. 2022 – Dec. 2023*
    - **Affiliation**: HKU NLP Group, Department of Computer Science, Faculty of Engineering
    - **Duty**: Conducting research on multimodal natural language processing and FinTech.

- **Microsoft (MSFT)** — Haidian, Beijing, China
  <u>*Research Intern*</u> *(Full-time, Onsite, Paid)* — *Nov. 2020 – Sep. 2021*
    - **Affiliation**: Natural Language Computing (NLC) Group, Microsoft Research Asia (MSRA)
    - **Duty**: Conducted research on multilingual neural machine translation.

PUBLICATIONS

Google Scholar ‖ Semantic Scholar ‖ OpenReview ‖ HuggingFace ‖ RG ‖ AMiner ‖ ORCID ‖ DBLP ‖ ACM ‖ ACL

Research Interests: Artificial Intelligence, Natural Language Processing, Multilingualism & Multimodality, World Model

UNDER REVIEW

- [P3] **UniCoder: Scaling Code Large Language Model via Universal Code**
  Tao Sun, Linzheng Chai, Jian Yang, <u>Yuwei Yin</u>, Hongcheng Guo, Jiaheng Liu, Bing Wang, Liqun Yang, Zhoujun Li
  Under review ‖ arXiv ‖ Code — *Mar 2024*

- [P2] **Red Teaming Visual Language Models**
  Mukai Li, Lei Li, <u>Yuwei Yin</u>, Masood Ahmed, Zhenguang Liu, Qi Liu
  OpenReview ‖ arXiv 2401.12915 ‖ RTVLM Data — *Jan 2024*

- [P1] **M³IT: A Large-scale Dataset towards Multimodal Multilingual Instruction Tuning**
  Lei Li, <u>Yuwei Yin</u>, Shicheng Li, Liang Chen, Peiyi Wang, Shuhuai Ren, Mukai Li, Yazheng Yang, Jingjing Xu, Xu Sun, Lingpeng Kong, Qi Liu
  OpenReview ‖ arXiv 2306.04387 ‖ M3IT Data — *Jun 2023*

CONFERENCE PROCEEDINGS

- [C8] **m³P: Towards Multimodal Multilingual Translation with Multimodal Prompt**
  Jian Yang, Hongcheng Guo, <u>Yuwei Yin</u>, Jiaqi Bai, Bing Wang, Jiaheng Liu, Xinnian Liang, Linzheng Chai, Liqun Yang, Zhoujun Li
  Accepted in Conference *LREC-COLING 2024* ‖ arXiv 2403.17556 ‖ InstrMulti102 Data — *May 2024*

- [C7] **GanLM: Encoder-Decoder Pre-training with an Auxiliary Discriminator**
  Jian Yang, Shuming Ma, Li Dong, Shaohan Huang, Haoyang Huang, <u>Yuwei Yin</u>, Dongdong Zhang, Liqun Yang, Zhoujun Li and Furu Wei
  Published in Conference *ACL 2023* ‖ Paper ‖ arXiv 2212.10218 ‖ Code — *Jul 2023*

- [C6] **HanoiT: Enhancing Context-aware Translation via Selective Context**
  Jian Yang, <u>Yuwei Yin</u>, Shuming Ma, Liqun Yang, Hongcheng Guo, Haoyang Huang, Dongdong Zhang, Yutao Zeng, Zhoujun Li, Furu Wei
  Published in Conference *DASFAA 2023* ‖ Paper ‖ arXiv 2301.06825 — *Apr 2023*

- [C5] **CROP: Zero-shot Cross-lingual Named Entity Recognition with Multilingual Labeled Sequence Translation**
  Jian Yang, Shaohan Huang, Shuming Ma, <u>Yuwei Yin</u>, Li Dong, Dongdong Zhang, Hongcheng Guo, Zhoujun Li, Furu Wei
  Published in Conference *EMNLP 2022 Findings* ‖ Paper ‖ arXiv 2210.07022 ‖ Code — *Dec 2022*

- [C4] **UM4: Unified Multilingual Multiple Teacher-Student Model for Zero-Resource Neural Machine Translation**
  Jian Yang*, Yuwei Yin*, Shuming Ma, Dongdong Zhang, Shuangzhi Wu, Hongcheng Guo, Zhoujun Li, Furu Wei
  Published in Conference *IJCAI-ECAI 2022* ‖ Paper ‖ arXiv 2207.04900 ‖ Code   *Jul 2022*
- [C3] **HLT-MT: High-resource Language-specific Training for Multilingual Neural Machine Translation**
  Jian Yang, Yuwei Yin, Shuming Ma, Dongdong Zhang, Zhoujun Li, Furu Wei
  Published in Conference *IJCAI-ECAI 2022* ‖ Paper ‖ arXiv 2207.04906 ‖ Code   *Jul 2022*
- [C2] **Improving Multilingual Neural Machine Translation with Auxiliary Source Languages**
  Weijia Xu, Yuwei Yin, Shuming Ma, Dongdong Zhang, Haoyang Huang
  Published in Conference *EMNLP 2021 Findings* ‖ Paper ‖ Code   *Nov 2021*
- [C1] **Multilingual Agreement for Multilingual Neural Machine Translation**
  Jian Yang, Yuwei Yin, Shuming Ma, Haoyang Huang, Dongdong Zhang, Zhoujun Li, Furu Wei
  Published in Conference *ACL-IJCNLP 2021* ‖ Paper ‖ Code   *Aug 2021*

JOURNAL ARTICLES
- [J1] **GTrans: Grouping and Fusing Transformer Layers for Neural Machine Translation**
  Jian Yang, Yuwei Yin, Shuming Ma, Haoyang Huang, Dongdong Zhang, Furu Wei, Zhoujun Li
  Published in Journal *IEEE TASLP* ‖ Paper ‖ arXiv 2207.14467 ‖ Code   *Nov 2022*

ARXIV PREPRINTS
- [A3] **MT4CrossOIE: Multi-stage Tuning for Cross-lingual Open Information Extraction**
  Zixiang Wang, Linzheng Chai, Jian Yang, Jiaqi Bai, Yuwei Yin, Jiaheng Liu, Hongcheng Guo, Tongliang Li, Liqun Yang, Hebboul Zine el-abidine, Zhoujun Li
  arXiv 2308.06552 ‖ Code   *Aug 2023*
- [A2] **FinPT: Financial Risk Prediction with Profile Tuning on Pretrained Foundation Models**
  Yuwei Yin, Yazheng Yang, Jian Yang, Qi Liu
  arXiv 2308.00065 ‖ Code ‖ FinBench Data   *Jul 2023*
- [A1] **TTIDA: Controllable Generative Data Augmentation via Text-to-Text and Text-to-Image Models**
  Yuwei Yin, Jean Kaddour, Xiang Zhang, Yixin Nie, Zhenguang Liu, Lingpeng Kong, Qi Liu
  arXiv 2304.08821 ‖ Code   *Apr 2023*

ACADEMIC SERVICE

REVIEWER (PROGRAM COMMITTEE)
- Conference Proceedings
  - ARR: ACL Rolling Review   *2023, 2024*
  - NAACL: Annual Conference of the North American Chapter of the Association for Computational Linguistics   *2024*
  - EACL: Conference of the European Chapter of the Association for Computational Linguistics   *2024*
  - ACL: Annual Meeting of the Association for Computational Linguistics   *2023*
  - EMNLP: Conference on Empirical Methods in Natural Language Processing   *2023*
  - ICML: International Conference on Machine Learning   *2023*
- Journal Articles
  - ncomms: Nature Communications   *2023*
  - TALLIP: ACM Transactions on Asian and Low-Resource Language Information Processing   *2023*
  - TASLP: IEEE/ACM Transactions on Audio, Speech, and Language Processing   *2022*

HONORS

SCHOLARSHIPS
- National Scholarship   (Awarded by The Ministry of Education of China)   *Nov. 2016 & Nov. 2017*
- Excellent Academic Scholarship   (Awarded by BISTU)   *Jan. 2016 & Jan. 2017 & Jan. 2018*
- Research and Innovative Scholarship   (Awarded by BISTU)   *Jan. 2016 & Jan. 2017*
- Social Contribution Scholarship   (Awarded by BISTU)   *Jan. 2017 & Jan. 2018*

AWARDS
- Award of Excellence in Stars of Tomorrow Internship Program   (Awarded by MSRA)   *Sep. 2021*
- Beijing Outstanding Graduate Award   (Awarded by Beijing Municipal Commission of Education)   *Jul. 2018*
- Outstanding Graduate Award   (Awarded by BISTU)   *Jul. 2018*
- Diligent Study & Research Pioneer Award   (Awarded by BISTU)   *Dec. 2017*

PRIZES
- 1st Prize in The 32nd National Physics Competition, Beijing District   (Awarded by Beijing Physics Society)   *Dec. 2015*
- 1st Prize in The 9th Mathematics Competition   (Awarded by BISTU)   *Jun. 2015*