

# XUANSHENG WU

Email: wuxsmail@163.com

Phone: +1 (213)-376-9286

Github: [www.github.com/JacksonWuxs](https://www.github.com/JacksonWuxs)

## EDUCATION

---

University of Georgia

08/2021 – Present

- Doctor of Philosophy in **Computer Science**

Shanghai University of International Business and Economics

09/2016 – 06/2020

- Bachelor of Science in **Applied Statistic**

## HONORS & AWARDS

---

- Rewarded **100+ citations** on Google Scholar 04/2024
- Rewarded **700+ Stars** on Github for first-author projects 03/2024
- Student Travel Award for attending WWW 2024 03/2024
- 2021 Tencent Advertising Algorithm Competition - **Top 1.7%** 07/2021
- Shanghai Aijian Scholarship - **Top 1%** 01/2020
- Research Pioneer Award of Shanghai University of International Business and Economics - **Top 1%** 08/2019
- 2nd Prize of China Programming Contest for College Students - **Top 4.5%** 08/2019
- Single-subject Scholarship of Shanghai University of International Business and Economics - **1/125** 11/2017
- 1<sup>st</sup> Asian University Archery Championship - **3rd** in Men's Group 04/2017

## PUBLICATIONS

---

- Understanding the Behavior Shift in LLMs after Instruction Tuning, **NAACL (Oral)**, 2024.
- Could Small Language Models Serve as Recommenders, **WWW (Oral)**, 2024.
- InFoBench: Evaluating Instruction Following Ability in Large Language Models, **ACL Finding**, 2024.
- Usable XAI: 10 Strategies Towards Exploiting Explainability in the LLM Era, ArXiv, 2024.
- Retrieval-Augmented In-context Model Editing for Multi-hop Question Answering, ArXiv, 2024.
- DIRECT: Dual Interpretable Recommendation with Multi-aspects Word Attribution, **TIST**, 2024.
- Black-box Backdoor Defense via Zero-shot Image Purification, **NIPS**, 2023.
- Matching Exemplar as Next Sentence Prediction (MeNSP), **AIED (Oral)**, 2023.
- NoPPA: Non-Parametric Pairwise Attention Random Walk Model for Sentence Representation, ArXiv, 2023.
- Rethinking the Impacts of Overfitting and Feature Quality on Small-scale Video Classification, **ACM MM (Oral)**, 2021.
- Lifestyle-based Approach for Cervical Cancer Screening, **ICDATA**, 2018.
- Optimization of Value Average Strategy in China Stock Market, **China Collective Economy**, 2018.

## INTERSHIPS

---

Samsung Research American

06/2024 – Present

- Proposed a series of local and global explanation methods for interpreting transformer-based language models.

Tencent – Tencent AI Lab

05/2023 – 08/2023

- Proposed a series of local and global explanation methods for interpreting transformer-based language models.
- Examined instruction tuning by comparing explanations generated by pre-trained and fine-tuned models. We found that: (1) instruction tuned models more effectively recognize instruction from user prompts to drive the response generation; (2) feed-forward networks and self-attention makes distinctive contributions to this process.

Baidu - Department of Natural Language Process

09/2020 – 04/2021

- Raised the Long Text Semantic Segmentation problem which leads to a significant negative impact of distorted paragraphs on downstream tasks (e.g., Machine Reading Comprehension, Query Generation, and ES Selection). Developed a document semantic segmenting model by designing training objectives and construction strategies with segmenting **F1=72.31%**. The model has been applied as a service to support Baidu Top-1 Search and FAQ Mining System.

- Investigated the bottleneck of the FAQ Mining System causing the low GPU utilization problem. Proposed a new architecture consisting of three components, namely Proxy, Scheduler, and Operator, which parallelly handles requests and schedules the GPU-based services dynamically according to the payload of each service. Designed a low-resource scheduling algorithm to assign services for different GPUs. Improved GPU Utilization by **2.9X** and QPS by **4.1X**.
- Worked with coworkers to improve searching experience of Baidu Top-1 Search under metrics **Good:Same:Bad=28:6:1** by optimizing services of Truncation Detection, Enumerate Answers, and Multi-Resources Recall.
- Deployed a query filter service for queries that require videos as answers (VQA) based on GBDT with **F1=93.43%**.

**Pingan OneConnect Co., Ltd - Institute of Big Data**

01/2019 – 06/2019

- Proposed to use Feature Context-Free Grammar for the Seq2SQL task, new solution supports 34 external patterns.

**Shanghai Ronghao Investment and Management Co., Ltd**

01/2018 – 02/2018

- Designed and tested a short-term trading strategy by analyzing over 40 millions transaction records.

## RESEARCHES

---

**Could Small Language Models Serve as System Cold-start Recommenders?**

09/2022 – 09/2023

- Formalized the system cold-start recommendation problem and launched the first benchmark for this challenge.
- Provided a mathematical framework of the in-context recommendation under the Hidden Markov assumption and demonstrated that in-context recommendation ability could be enhanced by model and prompt pre-training.
- Proposed a corpus refinement method and a decoupled prompt pretraining method to enhance small language models, empowered BERT-mini (11M parameters) achieves comparable performance with BERT-large (330M parameters).

**DIRECT: Dual Interpretable Recommendation with Multi-aspect Word Attribution**

04/2022 – 08/2022

- Proposed a review-based interpretable recommender system, predicting user preferences by averaging sentiment polarities of words weighted by word importance, where a word is important if it corresponds to an aspect of the item.
- Employed a concept-bottleneck layer and maximized the coding rate reduction on the space of aspect representations by leveraging a word-word affinity graph extracted from a pre-trained language model to learn discriminative aspects.
- Quantified experiments and case studies showed that DIRECT is comparable to SOTAs but provides clear explanations.

**NoPPA: Non-Parametric Pairwise Attention Random Walk Model for Sentence Representation**

09/2021 – 11/2021

- Proposed a novel sentence encoder by integrating the non-parametric self-attention into the bag-of-words model.
- Applied a kernel function to construct contextual word embeddings based on both positional and semantic embeddings.
- Evaluated NoPPA on eight different tasks and exceeded the bag-of-words-based baselines by **Acc=2.89%** on average.

**2021 Tencent Advertising Algorithm Competition: Multimodal Video Ads Tagging**

05/2021 – 07/2021

- Investigated that the baseline InceptionV3+Vggish+BERT+Resnet50-->NextVLAD-->ContextGate-->MLP has two issues: overfitting and outdated feature extractor, used the following solutions to improve it from 76.10% to **GAP=82.10%**.
- Removed the center frame modal as a redundant modality; tuned dropout rate; applied data augment strategies.
- Upgraded feature extractor; Introduced Temporal Shift Model; Fused ASR captions within image flow.

**Lifestyle-Based Cervical Cancer Screening**

03/2018 – 03/2019

- Improved the data engineering pipeline for feature selection and missing value handling, outperformed SOTA by 5%.

## PROJECTS

---

**Reproduction of Forward-Forward Algorithm**

01/2023

- Reproduced the Forward-Forward algorithm proposed by Geoffrey Hinton, rewarded **140+** stars on Github.

**DaPy: An open-source and easy-to-use data analysis framework for humans**

09/2017 – Present

- Achieved comparable efficiency with Pandas by using MemoryView, Cache-Friendly Operations, and Binary Search Index.

**Distributed Archery Events Supporting System based on B/S+C/S Hybrid Architecture**

12/2017 – 12/2018

- Proposed “B/S + C/S Hybrid Layout Architecture” to support high efficiency and high stability concurrently.
- Served 2018 Chinese University Archery Championships (peak of the system was **1430** visits per second).

## SKILLS & HOBBIES

---

- Python, C, Linux, SQL, SPSS, MATLAB.
- Archery (part-time coach 2+ years), Basketball (core member in school team), Travelling w/ my girlfriend and family.