Wei-Jie Xu

Nanjing University| xuwj@lamda.nju.edu.cn
https://vejaxu.github.io/

EDUCATION

Soochow University **Suzhou, Jiangsu**
Bachelor of Engineering in Artificial Intelligence Sep 2020 - June 2024

- *GPA*: 3.5 (Weighted Average Grade: 83.34)
- *Courses:* Machine Learning, Deep Learning, Introduction to Artificial Intelligence
- *Skills:* Programming languages (Python, C++), Operating system (Linux), Deep learning frameworks (PyTorch)
- *Bachelor's Thesis:* Design of a Long Document Knowledge Question-Answering System based on Large Language Models, using the LangChain framework.

Nanjing University **Nanjing, Jiangsu**
Master of Engineering in Computer Science and Technology Sep 2024 - June 2027

ACDEMIC EXPERIENCE

Currently, I'm interested in Machine Learning and Data Mining, especially in Anomaly Detection and Clustering. And I'm also interested in Large Language Models.

Spectral Clustering Sep 2024 – Nov 2024
Background:

- We propose a novel distribution-based spectral clustering which constructs a smaller bipartite graph between n points and k distributions, enabling the eigen-decomposition of only a k*k matrix while preserving clustering quality. Extensive experiments performed on synthetic and real-world datasets demonstrate the superiority and effectiveness of the proposed method compared to the state-of-the-art algotithms

Core Outputs:

- Is k * k Matrix Eigen-Decomposition Sufficient for Spectral Clustering ?

Structured Pruning of Large Language Models Nov 2024 - Jan 2025
Background:

- We propose, an innovative and effective pruning method that targets unimportant layers by leveraging the gradient discrepancy between perturbed and unperturbed states. Our comprehensive experimental evaluation, which spans a diverse array of LLMs across various series and scales, demonstrates that our method minimizes performance degradation post-pruning and outperforms existing state-of-the-art (SOTA) structured pruning methods under equivalent pruning ratios.

PROJECT EXPERIENCE

Huawei Ascend Intelligence Project Nov 2022 – June 2023
Background:

- *Focusing on MindSpore code migration, collaborated in a group led by the teacher and Huawei experts to migrate the OptInter paper from PyTorch to MindSpore.*

Core Outputs:

- Successfully reproduced and implemented the PyTorch version code, meeting Huawei's accuracy requirements (faced challenges in migrating to MindSpore due to differences in underlying logic, resulting in an unexpected project conclusion).

Bachelor's Thesis Feb 2024 – Jun 2024

Design of a Long Document Knowledge Question-Answering System based on Large Language Models, using the LangChain framework.

Campus & Professional Services

Campus Experience:

Assistant at the Student Work, School of Computer Science and Technology, Soochow University.

Assistant at the Graduate Student Association, School of Artificial Intelligence, Nanjing University.

Professional:

Currently None.

Scholarships & Awards

Scholarships:

New Point Software Scholarship, School of Computer Science and Technology, Soochow University, 2022.

Postgraduate Third Prize Academic Scholarship, Nanjing University, 2024.